

09/6/13 486

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein: - protein search, using sw model

Run on: November 7, 2003, 14:47:43 ; Search time 21 Seconds  
(without alignments)  
398.931 Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 991  
Sequence: 1 MELMSDSNLSNLVITDASSL.....GGVNTVPVSNLRQLGRREVM 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 423.0858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*

1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/6C\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/6D\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY-ES

result No.	Score	Query Match	Length	ID	Description
1	99.1	100.0	198	3	US-09-080-983-15
2	85.6	8.6	593	2	US-08-910-551B-2
3	8.2	732	4	US-09-252-991A-30074	Sequence 2, Appl
4	79.5	8.0	770	1	US-08-445-135-2
5	79.5	8.0	477	2	US-08-770-544-16
6	77.5	7.8	410	4	US-09-252-991A-20306
7	77.5	7.8	410	4	US-09-252-991A-20306
8	77.5	7.8	537	3	US-09-252-991A-32657
9	76.7	7.7	223	3	US-09-080-983-13
10	74.5	7.5	587	4	US-09-107-532A-6435
11	73.5	7.4	706	1	US-08-484-105-16
12	73.5	7.4	706	1	US-08-484-105-16
13	73.5	7.4	503	3	US-09-357-251-36
14	72.5	7.3	641	3	US-09-422-869-26
15	72.5	7.3	553	6	5310678-1
16	72.5	7.3	707	2	US-08-576-165-4
17	71.5	7.2	486	4	US-09-252-991A-31879
18	71.5	7.2	597	4	US-09-252-991A-21752
19	71.5	7.2	1048	4	US-09-171-699-10
20	71.5	7.2	193	4	US-08-671-548C-46
21	71.5	7.2	1040	4	US-09-328-352-7239
22	70.5	7.1	359	1	US-08-137-627-4
23	70.5	7.1	359	2	US-08-865-348-4
24	70.5	7.1	435	2	US-08-531-439B-4
25	70.5	7.1	586	4	US-09-107-532A-6918
26	69.5	7.0	206	4	US-09-107-532A-6952
27	69.5	7.0	414	4	US-09-252-991A-17176

Sequence 24, Appl  
Sequence 24, Appl  
Sequence 6, Appl  
Sequence 2, Appl  
Sequence 13, Appl  
Sequence 4807, Ap  
Sequence 6865, Ap  
Sequence 11, Appl  
Sequence 11, Appl  
Sequence 11, Appl  
Sequence 11, Appl  
Sequence 5, Appl  
Sequence 2, Appl  
Sequence 4377, Ap  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-09-080-983-15  
; Sequence 15, Application: US/09080983  
; Patent No. 6197948  
; GENERAL INFORMATION:  
; APPLICANT: Zhu, Hai-Ying  
; APPLICANT: Ling, Kai-Shu  
; APPLICANT: Gonsalves, Dennis  
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS  
; TITLE OF INVENTION: AND THEIR USES  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/080,983  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/047,194  
; FILING DATE: 20-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1631  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-080-983-15

Query Match: 100.0%; Score 991; DB 3; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1.7e-105;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 XELMSDSNLSNLVITDASSLNGVCKLLSAEVEKMLVQKAGNEGIEVVFGLLYALAAR 60  
DB 1 MEUMSDSNLSNLVITDASSLNGVCKLLSAEVEKMLVQKAGNEGIEVVFGLLYALAAR 60  
QY 61 TTSPKQVQADSDVIFSNSFGERNVVVTEGDLKKVLDGCAPLFRFTNKLTFRGTTEAYV 120  
DB 61 TTSPKQVQADSDVIFSNSFGERNVVVTEGDLKKVLDGCAPLFRFTNKLTFRGTTEAYV 120  
QY 121 DFCIAYKHLPOLNAAAEIGIPAEDSYLAADFLGCPKLSELQQRKMPASMYALKTEGG 180  
DB 121 DFCIAYKHLPOLNAAAEIGIPAEDSYLAADFLGCPKLSELQQRKMPASMYALKTEGG 180  
QY 181 VVNTPVSNLRQLGRREV 198  
DB 181 VVNTPVSNLRQLGRREV 198

## RESULT 2

US-08-910-551B-2  
Sequence 2, Application US/08910551B  
Patent No. 5910571  
GENERAL INFORMATION:  
APPLICANT: Sylvain Moineau, Barbara  
APPLICANT: J. Holler, Peter A. Vandenberg,  
APPLICANT: Ebenezer R. Vedamuthu, Jeffrey K.  
APPLICANT: Kondo  
TITLE OF INVENTION: DNA Encoding Phage  
TITLE OF INVENTION: Abortive Infection Protein  
TITLE OF INVENTION: From Lactococcus  
TITLE OF INVENTION: Lactis, and Method of Use Thereof  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Ian C. McLeod  
STREET: 2190 Commons Parkway  
CITY: Okemos  
STATE: Michigan  
COUNTRY: USA  
ZIP: 48864  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette 5.25 inch,  
MEDIUM TYPE: 360 Kb storage  
COMPUTER: Acer  
OPERATING SYSTEM: MS-DOS (version: 4)  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/310,551B  
FILING DATE: August 13, 1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/565,907  
FILING DATE: December 1, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ian C. McLeod  
REGISTRATION NUMBER: 20,931  
REFERENCE/DOCKET NUMBER: Quest 4.1-158  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (517) 347-4100  
TELEFAX: (517) 347-4103  
TELEX: No. 591057-e  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 599  
TYPE: Amino Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Protein  
DESCRIPTION: No  
HYPOTHETICAL: No  
ANTI-SENSE: No  
FRAGMENT TYPE: N/A  
ORIGINAL SOURCE:  
ORGANISM: Lactococcus lactis  
STRAIN:

INDIVIDUAL ISOLATE: W1  
DEVELOPMENTAL STAGE: N/A  
HAPLOTYPE: N/A  
TISSUE TYPE: N/A  
CELL TYPE: bacterium  
CELL LINE: N/A  
ORGANELLE: N/A  
IMMEDIATE SOURCE:  
LIBRARY: genomic  
CLONE: SMQ-20  
POSITION IN GENOME: N/A  
FEATURE:  
NAME/KEY: phage abortive infection  
LOCATION: N/A  
IDENTIFICATION METHOD: based upon DNA  
OTHER INFORMATION: phage resistance  
OTHER INFORMATION: enzyme Ab-E  
PUBLICATION INFORMATION: N/A  
US-08-910-551B-2

Query Match 8.8%; Score 85; DB 2; Length 599;  
Best Local Similarity 24.3%; Pred. No. 0.5;  
Matches 36; Conservative 26; Mismatches 46; Indels 40; Gaps 6;  
QY 2 ELMSDSNL-----SNLVTD-----ASSLNGVCK--KLSA 30  
DB 261 EFLNEFLICRENNLIINDNKTKVDNPPFVDSKSKSDIFSFFENITSTNSCKWKEISN 320  
QY 31 EVELMLVQKAGNEGIEVVFGLLYALAARTTSPKQVQADSDVIFSNSFGERNVVVTEG 89  
DB 321 FIDYCVNEEHLGNKAIGKIFPVI-----TNTLKQKVKDTKNI-DN-PSKKNVVTNFN 372  
QY 90 DLKKVLDGCAPLFRFTNKLTFRGTTE 117  
DB 373 VFELDLSLKOSRLTNKLTFFENINE 400

## RESULT 3

US-09-252-991A-30074  
Sequence 30074, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 30074  
LENGTH: 732  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30074  
Query Match 8.2%; Score 81; DB 4; Length 732;  
Best Local Similarity 22.4%; Pred. No. 1.9;  
Matches 30; Conservative 26; Mismatches 50; Indels 28; Gaps 5;  
QY 17 ASSLNGVCKLLSAEVEKMLVQKAGNEGIEVVFGLLYALAARTTSPKQVQADSDVIFS 76  
DB 44 ANTMGVYREAMKTIARLEAEK-----EGVL-----TSAKKTFAGGDL--- 87  
QY 77 NSFGERNVVVTEGDLKKVLDGCAPLFRFTNKLTFRGTTEAY-----VDFCIAYKH 128  
DB 88 -----NELIKVTADAPAFYQIGLELKGQLRRLETGLKPVVAAINGAALGGGWEICLACHH 143  
QY 129 KLPOLNAAAEIGIP 142

Db 144 RIALDNGVQLGLP 157

## RESULT 4

US-08-445-135-2  
; Sequence 2, Application US/08445135  
; Patent No. 5658789  
; GENERAL INFORMATION:  
; APPLICANT: Quaranta, Vito  
; APPLICANT: Hornia, Marketta  
; TITLE OF INVENTION: Promotion of Epithelial Cell Adhesion  
; TITLE OF INVENTION: and Hemidesmosome Assembly by a Laminin-Like Molecule  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,135  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/151,134  
; FILING DATE: 12-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israel, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: DESXOS.002CP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 770 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; IMMEDIATE SOURCE:  
; CLONE: 150 kD  
US-08-445-135-2

Query Match 8.0%; Score 79.5; DB 1; Length 770;

Best Local Similarity 26.6%; Pred. No. 3.1;  
Matches 38; Conservative 22; Mismatches 54; Indels 29; Gaps 7;

Qy 2 ELMSDLSNLVITDA--SSINGVDKLLSAEVEKMLVQKAPNEGIEVVGGLLVALAA 59

Db 325 ELINEAKWTKLQCEINPALNSLOQTLKTVSVQKLLDAN-----VTAV 369

Qy 60 RTTSPKVQADSDVIFNSGGERNVV-VTEGDLKKYLDGQCAPJTRFTNKLRT-FGRT--- 114

Db 370 RNDLRGIQRGIDISWVS---GAKSMVRKANGITSEVLDSAPSRRRIWEGIRATGCTRHE 426

Qy 115 -FTEAYVDFCIAYK---HKLPOL 133

Db 427 DFNKALIDANNSVKLTKXQPL 449

## RESULT 5

US-08-770-544-16  
; Sequence 16, Application US/08770544

; Patent No. 5907085  
; GENERAL INFORMATION:  
; APPLICANT: Gonsalves, Dennis  
; APPLICANT: Ling, Kai-Shu  
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND  
; TITLE OF INVENTION: THEIR USES  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle L.P  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/770,544  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60009008  
; FILING DATE: 21-DEC-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/621  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 477 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-770-544-16

Query Match 8.0%; Score 79; DB 2; Length 477;

Best Local Similarity 21.0%; Pred. No. 1.8;  
Matches 43; Conservative 40; Mismatches 60; Indels 62; Gaps 11;

Qy 12 LVITDASSLNGVDKLLSAEVEKMLVQK---APNE-GIEVVGGLLYALAAATTSFKVQ 67

Db 302 LPITTEALQIN-----ARLRKLVLSKSGSQTPROXGNMIVAMIQFLVLYSTVKXISVK 353

Qy 68 ---RADSDVIFNSFGERNVVVTGGDLKKYL---DGCAPLRTFTNKLRTFTFTTAYV 120

Db 354 DGYRVETEL-----GQKRVLYSYSEVREAILGKYGASP-----TNTVRSFMYFAHTT 403

Qy 121 DFCIAYKHKLQPLNAAELGIP-----AEDSYLAADFGLGTCPKLSELQOSRXM 168

Db 454 TLLIEKKIQ-PACTALAKHGVPKRFPTPCDFALLDNRYYPADVLT-----KANA 451

Qy 169 FASMYALKTEGVVNTFVSNLRQLG 193

Db 452 MACAIAIKS-----ANLRKKG 467

## RESULT 6

US-09-579-259-16  
; Sequence 16, Application US/09579259  
; Patent No. 6558953

; GENERAL INFORMATION:  
; APPLICANT: Gonsalves, Dennis

; APPLICANT: Ling, Kai-Shu

; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS  
; PROTEINS AND THEIR USES  
; NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14603

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICANT: Patent In Release #1.0, Version #1.30  
FILING DATE: 25-May-2000  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION NUMBER: US 6009008

FILING DATE: 21-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 19603/621

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 263-1304

TELEFAX: (716) 263-1603

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 477 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 16:

US-09-579-259-16

Query Match 8.0%; Score 79; DB 4; Length 477;  
Best Local Similarity 21.0%; Pred. No. 1.8;  
Matches 43; Conservative 40; Mismatches 60; Indels 62; Gaps 11;

QY 12 LVITDASSLVGVKLLSAEVEKMLVQKQ---APNE-GIEVVFGLLLVALAARTTSPKQV 67

Db 302 LPITEALQIN-----AELRLVLSKSGSOTPRDMGNKIVANIQLFVLYSTVKNISVK 353

QY 68 ---RADSDVFSNGFGRNVTVEGDKKVL---DGCAPLTFRTNKLRTFGRTTEAVV 120

Db 354 DGYRVETEL-----GCKXVY:SYSEVREAILGKYGASP-----TNTVRSFNRYPAHTI 403

QY 121 DFCIAYKHLKPLQNAAEELGIP-----AECSYLAADFGLGTCFKLSELOQSRKM 168

Db 404 TLLIEKKIG-PACTALAKFGVPRKFTFYCFDFALLDNRYYPADVL-----KANA 451

QY 169 FASMAVAKTEGGVNTVPSNLRQLG 193

Db 452 MACATAIKS-----ANLRKXG 467

RESULT 7  
US-09-252-991A-20306  
; Sequence 20306, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 07196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/374,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/394,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142

Query Match 7.8%; Score 77; DB 4; Length 410;  
Best Local Similarity 25.0%; Pred. No. 2.1;  
Matches 60; Conservative 31; Mismatches 90; Indels 59; Gaps 13;

QY 7 SNLSNLVITDASSLVGVKLLSAEVEKMLVQKQ-----GAPNEGIEVV 49

Db 78 SCMDGLAFURHASLSGKSHVLSSEVDPLRQATISMEICLGLNFGDLGKPFSEJER 136

QY 50 FGLLYVALAARTTSPK-VQRAD-----SOVIFSNFSGE-----RNVVTEGDL--KKVLD 96

Db 137 TALLTRYNNARODLPQIEVAELPSVADVVRGVDNGEFEAYYQPKVALDGGGLIGAEVLA 196

; SEQ ID NO 20306

; LENGTH: 410

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-20306

Query Match 7.8%; Score 77.5; DB 4; Length 410;  
Best Local Similarity 25.0%; Pred. No. 2.1;  
Matches 60; Conservative 31; Mismatches 90; Indels 59; Gaps 13;

QY 7 SNLSNLVITDASSLVGVKLLSAEVEKMLVQKQ-----GAPNEGIEVV 49

Db 78 SCMDGLAFURHASLSGKSHVLSSEVDPLRQATISMEICLGLNFGDLGKPFSEJER 136

QY 50 FGLLYVALAARTTSPK-VQRAD-----SOVIFSNFSGE-----RNVVTEGDL--KKVLD 96

Db 137 TALLTRYNNARODLPQIEVAELPSVADVVRGVDNGEFEAYYQPKVALDGGGLIGAEVLA 196

QY 97 -----GCAPLTFRTNKLRT-----FGRTTEAVVDFCAYKHLKPLQNAAEELGIP 142

Db 197 RWRHPHLGVPPSHFLYVMETYNLVKLFWQFSQG-----LARRKLAQLGQPINLAFN 251

QY 143 AEDSYLAADF--GTCPKLSE--LQSRKMFASNYALKTEGGVNTVPSNLRQLGRREV 198

Db 252 VHPQLGSRALAEINISALLTEFHLPPSSVMFEI-----TETGLISAPASSLENLVRLRM 316

RESULT 8

US-09-252-991A-32657

; Sequence 32657, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/394,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 32657

; LENGTH: 597

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-32657

Query Match 7.8%; Score 77; DB 4; Length 597;  
Best Local Similarity 34.0%; Pred. No. 4.1;  
Matches 33; Conservative 8; Mismatches 44; Indels 12; Gaps 4;

QY 99 APLTRFTNKLRTFGRTTEAVVDFC-IAYKHLKPLQNAAEELGIPAEEDSYLAADF-GTCTP 157

Db 461 AALFRGTTELLHPSGSTILEADILCVIGHEDLP---ALGKLFSSQAPDRGLGARFFGDFV 517

QY 158 KLSELOQSRKMFASNYALKTEGGVNTVPSNLRQLGR 194

Db 519 LEGDAQLS--AVASLYGLKLDG-----IDGEOALGR 546

RESULT 9

US-09-080-983-13

; Sequence 13, Application US/09080983

; Patent No. 6197948

; GENERAL INFORMATION:

; APPLICANT: Zhu, Hai-Ying

; APPLICANT: Ling, Kai-Shu

; APPLICANT: Gonsalves, Dennis

; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS

; NUMBER OF SEQUENCES: 23



; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent'n Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/680,983  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/047,194  
; FILING DATE: 20-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/163:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 223 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-080-983-13

Query Match 7.7%; Score 76; DB 3; Length 223;  
Best Local Similarity 23.3%; Pred. No. 1.3;  
Matches 48; Conservative 36; Mismatches 88; Indels 34; Gaps 11;

QY 4 MDSNLSNLVITDASSINGVD-KLLSNAEVEKMLVQKAPKEGIEVVGGLLYAL--AAR 60  
DB 36 LKQENYSSV---DSSLSDSEVKEVLEKSKSPKSELASTDE--HFVYHIFFLIRCAKI 90  
QY 61 TTSKPVQRAEDSDVIFSNFSGERNVVVTEGJLKKVLDG-----CAPLT---RFTNKLRTFG 112  
DB 91 STSEKVKYVGS-----H-YVVDGKTYTLDWVFNMMKSLTKYKRVNGLRAFC 139  
QY 113 RTFTAYVDYFCIAYKHKLPOLNAAAEGLIPAEDSYLADEF--GTCPKLSLQSQSRKMFAS 172  
DB 140 CACEDLYLTVPAPINSERF-KTKAVGMKGLPVGKEYLGADEFSLGTSKLSMCHDRAVSIVAA 198  
QY 172 MYALKTE---GGVNTVPVSNLRQLGR 194  
DB 199 KNAVDRSAFTGG--ERKIVSLYDLGR 222

RESULT 10  
US-09-107-532A-6405  
; Sequence 6405, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM, ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 6405:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 587 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...587  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6405:  
US-09-107-532A-6405

Query Match 7.5%; Score 74.5; DB 4; Length 587;  
Best Local Similarity 22.6%; Pred. No. 7.8;  
Matches 43; Conservative 27; Mismatches 65; Indels 55; Gaps 10;  
QY 1 MELASDSSLNLVITDASS---LNGVDEKXLLSARVEKXVCXGAPNEG-EVVF-----50  
DB 224 MEESGKLNHLXNAENKTETFTIGIKXKQKQEVAKLVE-----NFMKQVFDGFFHA 278  
QY 51 ---GLLYALAAATTSKPVQRAEDSDVIFSNFSGE---RNWVTEGDLKKVLDGCAPLTR 103  
DB 279 DHPGNLLFHLVLTKEEQTCASR-KTETVHEKEFGSPAFRASTSAE-----DPVAP---327  
QY 104 FTNKLRTFGRTFTAYVDF-----CIAYKHKLPOLNAAAEGLIPAEDSY-LAADELSTC 156  
DB 328 -----YTINYIDFGWGHLSAGLRQKLTQ-----AVLALYTKDAYRIEKAVALRLC 372  
QY 157 PKLSLQSQSR 166  
DB 373 QQEGSFDES 382

RESULT 11  
US-08-484-105-16  
; Sequence 16, Application US/08484105  
; Patent No. 5589341  
; GENERAL INFORMATION:  
; APPLICANT: STILLMAN, Bruce  
; APPLICANT: BELL, Stephen P  
; APPLICANT: KOBAYASHI, Ryuji  
; APPLICANT: RINE, Jasper  
; APPLICANT: FOSS, Margit  
; APPLICANT: McNALLY, Francis J  
; APPLICANT: LAURENSEN, Patricia  
; APPLICANT: HERSKOWITZ, Ira  
; APPLICANT: Li, Joachim J  
; APPLICANT: GAVIN, Kimberly  
; TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DCS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,105  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman Ph.D., Richard Aron  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 706 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-105-16

Query Match 7.4%; Score 73.5; DB 1; Length 706;  
Best Local Similarity 19.3%; Pred. No. 13;  
Matches 43; Conservative 33; Mismatches 68; Indels 79; Gaps 9;

QY 10 SNLVITDASSINGVDKLLSAEVEKMLVQKGAPE-----GIVVFGLLLYALAARTTSP 64  
DB 479 SRLIVAVANTMDLPERILSNRISRLGLSRVFPYTHQTLEII-----IAARLEAV 531  
QY 65 KVRADSDVIFSNS--FGERNVVVTEGDLKKVLDGC----- 98  
DB 532 R-----DDVFSDDA:RFAARKVAASGDARRALD:ICRRASELAENKNGKVTPLIHQAIS 587  
QY 99 -----APLTR-----FTNKLRTFGRTTEAYVDFCIAYKHKLPOLNAAEL 139  
DB 588 EMVTSPLQKVLRLNLSFMQKVFCLCAIVNMRMSG--FAESVYVEVLEAEERLLRVMTTPD- 644  
QY 140 GIPAEYSYLAADFLGTCPKLSLQSRKMFASVYALKTEGGVV 192  
DB 645 ---AE-----AKFGELILRRPFGYVLSLSENGVL 672

RESULT 12  
US-08-484-106-16  
Sequence 16, Application US/08484106  
Patent No. 5614618  
GENERAL INFORMATION:  
APPLICANT: STILLMAN, Bruce  
APPLICANT: BELL, Stephen P  
APPLICANT: KOBAYASHI, Ryuji  
APPLICANT: RINE, Casper  
APPLICANT: POSS, Margit  
APPLICANT: McNALLY, Francis J  
APPLICANT: LAURENSEN, Patricia  
APPLICANT: HERSKOWITZ, Ira  
APPLICANT: LI, Joachim J  
APPLICANT: GAVIN, Kimberly  
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DCS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,106  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman Ph.D., Richard Aron  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 706 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-106-16

Query Match 7.4%; Score 73.5; DB 1; Length 706;  
Best Local Similarity 19.3%; Pred. No. 13;  
Matches 43; Conservative 33; Mismatches 68; Indels 79; Gaps 9;

QY 10 SNLVITDASSINGVDKLLSAEVEKMLVQKGAPE-----GIVVFGLLLYALAARTTSP 64  
DB 479 SRLIVAVANTMDLPERILSNRISRLGLSRVFPYTHQTLEII-----IAARLEAV 531  
QY 65 KVRADSDVIFSNS--FGERNVVVTEGDLKKVLDGC----- 98  
DB 532 R-----DDVFSDDA:RFAARKVAASGDARRALD:ICRRASELAENKNGKVTPLIHQAIS 587  
QY 99 -----APLTR-----FTNKLRTFGRTTEAYVDFCIAYKHKLPOLNAAEL 139  
DB 588 EMVTSPLQKVLRLNLSFMQKVFCLCAIVNMRMSG--FAESVYVEVLEAEERLLRVMTTPD- 644  
QY 140 GIPAEYSYLAADFLGTCPKLSLQSRKMFASVYALKTEGGVV 192  
DB 645 ---AE-----AKFGELILRRPFGYVLSLSENGVL 672

RESULT 13  
US-09-357-251-36  
Sequence 36, Application US/09357251  
Patent No. 6271441  
GENERAL INFORMATION:  
APPLICANT: Falco, S. Carl  
APPLICANT: Pamodu, Sayo O.  
APPLICANT: Orozco, Buddy  
APPLICANT: Schwaber, James S.  
TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase  
FILE REFERENCE: BB-1193  
CURRENT APPLICATION NUMBER: US/09/357,251  
CURRENT FILING DATE: 1999-07-20  
EARLIER APPLICATION NUMBER: 60/093,530  
EARLIER FILING DATE: July 21, 1998  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 36  
LENGTH: 503  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-357-251-36

Query Match 7.4%; Score 73; DB 3; Length 503;  
Best Local Similarity 24.6%; Pred. No. 9.2;  
Matches 48; Conservative 33; Mismatches 74; Indels 40; Gaps 12;  
  
QY 4 MDSNLSNLV-ITDASSLNGVDKLLSAEVEKMLVCKG-----APNEGIEVVFESJLLYAL 57  
DB 134 LTDTQSLIAQIKNSHLSIDAKILNDJKKRKLROGKITDFSVTGKPE--FSTDLTKL 191  
QY 58 AARTSPKQV-RADSDVIFS-NSFGERNVVVTEGDLKVLGCGAPLRTFNKLT--FGR 113  
DB 192 ETCJTSQVSTNAYKDLKFPYFNFSQGVQISSGALH-----PLNKVREERQIFFSM 244  
QY 114 TFE-----AYVD-----FCIAKXKLPOLNAAELGIPAEDSVLAADFLCTCPKLSLQ 164  
DB 245 GFTMPNQVYETGFWFDALY---VPOQHPARDL-----QDTFYKD-----ELTAEJFD 292  
QY 165 SRKYFASMYALKTEG 179  
DB 293 DKYMONIKAVHEQG 307

RESULT 14  
US-09-422-869-26  
Sequence 26, Application US/09422869  
Patent No. 6235481  
GENERAL INFORMATION:  
APPLICANT: POLONSKI, KENNETH S.  
APPLICANT: HORIKAWA, YUKIO  
APPLICANT: ODA, NACHISA  
APPLICANT: COX, NANCY J.  
APPLICANT: SREEMAN, SEAMUS  
APPLICANT: ZHOC, YUN-PING  
APPLICANT: OTANI, KENICHI  
APPLICANT: HANIS, CRAIG L.  
APPLICANT: BELL, GRAEME I.  
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES  
FILE REFERENCE: ARCD:307  
CURRENT APPLICATION NUMBER: US/09/422,869  
CURRENT FILING DATE: 1999-10-21  
EARLIER APPLICATION NUMBER: 60/134,175  
EARLIER FILING DATE: 1999-05-13  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 26  
LENGTH: 641  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-422-869-26

Query Match 7.3%; Score 72.5; DB 3; Length 641;  
Best Local Similarity 25.4%; Pred. No. 15;  
Matches 29; Conservative 19; Mismatches 43; Indels 23; Gaps 6;  
  
QY 79 FGERNVVTEGDLKVLGCGAPLRTFNKLTFTGRTFE-AYVDFCIAYKHKLPLNAAA 137  
DB 128 FGEWTEVVID-DLLPTINGDLVFS-FSTSMNEFNALLEKAYAK-LGQYE-----A 176  
QY 138 ELGIPAEDSVLAADFLCTCPKLSLQSR-----KMFASMYALKTEG 182  
DB 177 LDGLTITD--IIMDFTGTGLAEIIMQGRYTDLVEEKYKLFGLYKFTTKGGLI 228

RESULT 15  
5310678-1  
Patent No. 5310678  
APPLICANT: Bingham, Richard W.; Chambers, Philip; Emerson, Peter  
T. Millar, Neil S.  
TITLE OF INVENTION: NEWCASTLE DISEASE VIRUS GENE CLONES  
NUMBER OF SEQUENCES: 3  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/438,945  
FILING DATE: 17-NOV-1989

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 885,765  
FILING DATE: 15-JUL-1986  
SEQ ID NO: 1  
LENGTH: 553  
5310678-1  
  
Query Match 7.3%; Score 72; DB 6; Length 553;  
Best Local Similarity 24.1%; Pred. No. 14;  
Matches 49; Conservative 28; Mismatches 94; Indels 32; Gaps 8;  
  
QY 3 LMDSNLSNLVITDASSLN-GVCKLLSAEVEKMLVCKGAPNEGIEVVF-----GLLYAL 57  
DB 289 LPSVGNLNNRATYJELSVSTTRGFASALVPKVTQVGSVIELDTSYCIETOLDLY-- 346  
QY 58 AARTSPKQVRAADSDVIFSNSFGERNVVV---TEGDLK-----KVLGCGAPLT-RFT 105  
DB 347 CTRIVTPFM---SPQIYSCLSGNTSACMYSKTEGALTTPYMTIKGSVIANCKMTTCRCV 402  
QY 106 NKLRTEGRTFTTEAYVDPCIAKXKLPOLNAAELGIPAEDSVLAADFLCTCPKLSLQCS 165  
DB 403 NPPGIISQNYGEAV-----SLDPKQSCNVLSDGITLRLSGEPDATYQKNISIQCS 453  
QY 166 KMFASMYALKTEG 188  
DB 454 QVIITGNLIDISTELGNVNSISN 476

Search completed: November 7, 2003, 14:50:55  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2003, 14:47:43 ; Search time 21 seconds  
(without alignments)  
398.93; Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 991

Sequence: 1 MELASDSNLSNLVITDASSL.....GGVVNTPVSNRLQLGRREV 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	198	3	US-09-080-983-15
2	85	8.6	599	2	US-08-910-551B-2
3	81	8.2	732	4	US-09-252-991A-30074
4	79.5	8.0	770	1	US-08-445-135-2
5	79	8.0	477	2	US-08-770-544-16
6	79	8.0	477	4	US-09-579-259-16
7	77.5	7.8	410	4	US-09-252-991A-20306
8	77	7.8	597	4	US-09-252-991A-32657
9	76	7.7	223	3	US-09-080-983-13
10	74.5	7.5	587	4	US-09-107-532A-6405
11	73.5	7.4	706	1	US-08-484-105-16
12	73.5	7.4	706	1	US-08-484-106-16
13	73	7.4	503	3	US-09-357-251-36
14	72.5	7.3	641	3	US-09-422-869-26
15	72	7.3	553	6	5310678-1
16	72	7.3	707	2	US-08-576-165-4
17	71.5	7.2	486	4	US-09-252-991A-31879
18	71.5	7.2	597	4	US-09-252-991A-21752
19	71.5	7.2	1048	4	US-09-171-699-10
20	71	7.2	193	4	US-08-671-548C-46
21	71	7.2	1040	4	US-09-328-352-7238
22	70.5	7.1	359	1	US-08-137-627-4
23	70.5	7.1	359	2	US-08-865-348-4
24	70.5	7.1	435	2	US-08-531-439B-4
25	70.5	7.1	586	4	US-09-107-532A-6918
26	69.5	7.0	206	4	US-09-107-532A-6952
27	69.5	7.0	414	4	US-09-252-991A-17176

28 69.5 7.0 1713 3 US-08-600-982-24 Sequence 24, Appl  
29 69.5 7.0 17-3 3 PCT-US94-10261A-24 Sequence 24, Appl  
30 69.5 7.0 2042 4 US-09-077-098A-6 Sequence 6, Appl  
31 69.5 7.0 2285 4 US-09-308-375-2 Sequence 2, Appl  
32 69 7.0 363 3 US-08-984-618-13 Sequence 13, Appl  
33 68.5 6.9 382 4 US-09-107-532A-4807 Sequence 4807, Ap  
34 68.5 6.9 398 4 US-09-328-352-6865 Sequence 6865, Ap  
35 68 6.9 693 2 US-08-463-620-11 Sequence 11, Appl  
36 68 6.9 693 2 US-08-463-620-11 Sequence 11, Appl  
37 68 6.9 693 2 US-08-914-853-11 Sequence 11, Appl  
38 68 6.9 693 5 PCT-US95-03934A-11 Sequence 11, Appl  
39 68 6.9 892 4 US-09-336-447A-5 Sequence 5, Appl  
40 68 6.9 911 1 US-08-596-585-2 Sequence 2, Appl  
41 67.5 6.8 424 4 US-09-328-352-4377 Sequence 4377, Ap  
42 67.5 6.8 875 1 US-08-460-547A-10 Sequence 10, Appl  
43 67.5 6.8 875 1 US-08-250-847B-10 Sequence 10, Appl  
44 67.5 6.8 875 2 US-08-463-949A-10 Sequence 10, Appl  
45 67.5 6.8 875 3 US-08-464-410A-10 Sequence 10, Appl

#### ALIGNMENTS

#### RESULT: 1

US-09-080-983-15  
; Sequence 15, Application US/09080983  
; Patent No. 6:97948  
; GENERAL INFORMATION:  
; APPLICANT: Zhu, Hai-Ying  
; APPLICANT: Ling, Kai-Shu  
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS  
; TITLE OF INVENTION: AND THEIR USES  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentir Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/080,983  
; FILING DATE:  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 60/047,194  
; FILING DATE: 20-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1631  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-080-983-15

Query Match 100.0%; Score 991; DB 3; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1.7e-105;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 MELMSDNLNLVITDASSLNGVDKLLSAEVEKMLVKGAPNEGIEVVFGLLYALAAR 60  
DB 1 MELMSDNLNLVITDASSLNGVDKLLSAEVEKMLVKGAPNEGIEVVFGLLYALAAR 60  
CY 61 TTSPPKQVADSDVIFSNFSGERNVVVTEGDLKKVLDGCAPLRTFTNKLRTGRTFTTAYV 120  
DB 61 TTSPPKQVADSDVIFSNFSGERNVVVTEGDLKKVLDGCAPLRTFTNKLRTGRTFTTAYV 120  
CY 121 DFCIAYKHKLPLNAAAELGIPABSYLAADF.GTCPKLSELQOSRKMPFASMYALKTEGG 180  
DB 121 DFCIAYKHKLPLNAAAELGIPABSYLAADF.LGTCPKLSLQOSRKMPFASMYALKTEGG 180  
CY 181 VNTPVSNLRQLGRREV 198  
DB 181 VNTPVSNLRQLGRREV 198

RESULT 2

US-08-910-551B-2  
; Sequence 2, Application US/0891055:B  
; Patent No. 5910571  
; GENERAL INFORMATION:  
; APPLICANT: Sylvain Meineau, Barbara  
; APPLICANT: J. Holler, Peter A. Vandenbergh,  
; APPLICANT: Ebenezer R. Vedamuthu, Jeffrey K.  
; APPLICANT: Kondo  
; TITLE OF INVENTION: DNA Encoding Phage  
; TITLE OF INVENTION: Abortive Infection Protein  
; TITLE OF INVENTION: From Lactococcus  
; TITLE OF INVENTION: Lactis, and Method of Use Thereof  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ian C. McLeod  
; STREET: 2190 Commons Parkway  
; CITY: Okemos  
; STATE: Michigan  
; COUNTRY: USA  
; ZIP: 48864  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette 5.25 inch,  
; MEDIUM TYPE: 360 Kb storage  
; COMPUTER: Acer  
; OPERATING SYSTEM: MS-DOS (version 4)  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/910,551B  
; FILING DATE: August 11, 1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/565,907  
; FILING DATE: December 1, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ian C. McLeod  
; REGISTRATION NUMBER: 20,931  
; REFERENCE/DOCKET NUMBER: Quest 4.1-158  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (517) 347-4100  
; TELEFAX: (517) 347-4103  
; TELE: No. 5910571e  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 599  
; TYPE: Amino Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
; MOLECULE TYPE:  
; DESCRIPTION: Protein  
; HYPOTHEICAL: No  
; ANTI-SENSE: No  
; FRAGMENT TYPE: N/A  
; ORIGINAL SOURCE:  
; ORGANISM: Lactococcus lactis  
; STRAIN:

INDIVIDUAL ISOLATE: W1  
DEVELOPMENTAL STAGE: N/A  
HAPLOTYPE: N/A  
TISSUE TYPE: N/A  
CELL TYPE: bacterium  
CELL LINE: N/A  
ORGANELLE: N/A  
IMMEDIATE SOURCE:  
; LIBRARY: genomic  
; CLONE: SMQ-20  
; POSITION IN GENOME: N/A  
; FEATURE:  
; NAME/KEY: phage abortive infection  
; LOCATION: N/A  
; IDENTIFICATION METHOD: based upon DNA  
; OTHER INFORMATION: phage resistance  
; OTHER INFORMATION: enzyme Ab1E  
; PUBLICATION INFORMATION: N/A  
US-08-910-551B-2  
Query Match 8.6%; Score 85; DB 2; Length 599;  
Best Local Similarity 24.3%; Pred. No. 0.5;  
Matches 36; Conservative 26; Mismatches 46; Indels 40; Gaps 6;  
CY 2 ELMSDNL----SNLVITD-----ASSLNGVDK--KLISA 30  
DB 261 EPLNEFNLCRENLIINDNKTVDNPFVDKSSKSDIFFSPENITSTNSNKKWKEISN 320  
CY 31 EVEKMLVKGAPNEG-IEVVFGLLYALAARTTSPKVCRAOSDVTFNSFSGERNVAVTEG 89  
DB 321 FIDYCVNEEHLGNKGAKCIPFVI-----TNTTKQKKVDTKNI-DNIFSKRNVNFN 372  
CY 90 D.LKKVLDGCAPLRTFTNKLRTGRTFTT 117  
DB 373 VFEKILDLSDKDSRLTNKFLTFENINE 480  
RESULT 3  
US-09-252-991A-30074  
; Sequence 30074, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-19  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30074  
; LENGTH: 732  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30074  
Query Match 8.2%; Score 81; DB 4; Length 732;  
Best Local Similarity 22.4%; Pred. No. 1.9;  
Matches 30; Conservative 26; Mismatches 50; Indels 28; Gaps 5;  
CY 17 ASSLNGVDKLLSAEVEKMLVKGAPNEGIEVVFGLLYALAARTTSPKVCRAOSDVFS 76  
DB 44 ANTMNGVYREMAKTARLEAEK---EGIA---GWV-----TSAKKTFAGGGL--- 87  
CY 77 NSFGERNVVVTGDLKKVLDGCAPLRTFTNKLRTGRTFTTAYV-----VDFCIAYKH 128  
DB 88 -----NEJIKVTKADAPAFYCGIILELKGQLRRLETIGKPVVAINGAALGGGWEICLCHH 143  
CY 129 KLPQLNAAAEELGIF 142

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2003, 14:49:54 ; Search time 23 Seconds

(without alignments)  
1172.633 Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 931

Sequence: 1 MELMSDSNLSNLVITDASSL.....GGVNTFVSNLRLGRREV 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2600000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications: AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pcp.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pcp.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pcp.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pcp.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pcp.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pcp.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pcp.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pcp.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pcp.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pcp.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pcp.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pcp.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pcp.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pcp.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pcp.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pcp.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pcp.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	286	28.9	204	12	US-10-138-842A-36
2	120.5	12.2	223	12	US-10-138-842A-37
3	79	8.0	451	9	US-09-815-242-10721
4	79	8.0	477	12	US-10-138-842A-16
5	79	8.0	477	12	US-10-039-112-16
6	77	7.8	696	12	US-10-032-585-7808
7	76.5	7.7	356	15	US-10-156-761-11789
8	75	7.6	2626	10	US-09-793-508B-2
9	74.5	7.5	292	15	US-10-156-761-10849
10	74	7.5	368	15	US-10-205-823-377
11	74	7.5	399	9	US-09-925-301-1385
12	73.5	7.4	451	15	US-10-156-761-12518
13	73.5	7.4	833	9	US-09-844-281-1
14	73	7.4	485	15	US-10-156-761-13439
15	73	7.4	1987	12	US-10-032-585-7518

Sequence 2, Appli.  
Sequence 26, Appli.  
Sequence 18, Appli.  
Sequence 7333, Ap  
Sequence 10, Appli  
Sequence 3, Appli  
Sequence 5013, Ap  
Sequence 10878, A  
Sequence 14, Appli  
Sequence 12853, A  
Sequence 5570, Ap  
Sequence 12280, A  
Sequence 6806, Ap  
Sequence 3, Appli  
Sequence 113, App  
Sequence 6, Appli  
Sequence 2, Appli  
Sequence 7, Appli  
Sequence 394, App  
Sequence 666, Ap  
Sequence 8805, Ap  
Sequence 7360, Ap  
Sequence 42, Appli  
Sequence 1013, Ap  
Sequence 1934, Ap  
Sequence 1, Appli  
Sequence 1934, Ap  
Sequence 1934, Ap  
Sequence 5, Appli

# ALIGNMENTS

RESULT 1  
US-10-138-842A-36  
; Sequence 36, Application US-10138842A  
; Publication No. US20030148390A1  
; GENERAL INFORMATION:  
; APPLICANT: GONSALVES, DENNIS  
; APPLICANT: LING, KAI-SHU  
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND  
; TITLE OF INVENTION: THEIR USES  
; FILE REFERENCE: 07678/025006  
; CURRENT APPLICATION NUMBER: US/10138,842A  
; CURRENT FILING DATE: 2002-05-03  
; PRIOR APPLICATION NUMBER: US 09/579,259  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 09/224,898  
; PRIOR FILING DATE: 1998-12-31  
; PRIOR APPLICATION NUMBER: US 08/770,544  
; PRIOR FILING DATE: 1996-12-20  
; PRIOR APPLICATION NUMBER: US 60/009,008  
; PRIOR FILING DATE: 1995-12-21  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 204  
; TYPE: PRT  
; ORGANISM: Beet Yellow Virus  
US-10-138-842A-36

Query Match 28.9%; Score 286; DB 12; Length 204;  
Best Local Similarity 37.6%; Pred No. 1.8e-24;  
Matches 71; Conservative 29; Mismatches 87; Indels 2; Gaps 2;  
Qy 7 SKLSNLVITDASSLNGVDPKLLLSAEVEKMLVQKAPNEGIEVVFGLLYALAAATTSPKV 66  
Db 11 ATFNVSADQ\*CLHGECDCDKLRKNFECCLKLGVPEDNLGLGCLYSCATIGTSNKV 70  
Qy 67 QRASDVIFSNFS-GERNVVVTEGDLKKVLOGCAPLTFSTKCLRTFGSTFTAEVDFCIA 125







```
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204C89
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11789
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11789

Query Match          7.7%  Score 76.5;  DB 15;  Length 356;
Best Local Similarity 30.1%;  Pred. No. 4.4;
Matches 34;  Conservative 12;  Mismatches 38;  Indels 29;  Gaps 6;

QY 57 LAARTSPKVQRADSDVIFSNFSGERNVVTEGDKKVLGGCAPLRTFTNKL-RTFGRTF 115
Db 239 LAARHQAQVTTGDPR-LYENG-PEQSQVAIRGCAEKADACAAIGRDVKELDKVLTGPF 297

QY 116 T-----EAYVDFCIAYKHKLQLNAAAEELGI-----PAEDSYLAAD 151
Db 298 TPDGRPLESLDAFDFV-AGRH-----MELGITDIAHWPIPDSDFAAD 340

RESULT 8
US-09-759-508B-2
; Sequence 2, Application US/09759508B
; Publication No. US20020182599A1
; GENERAL INFORMATION:
; APPLICANT: Fishman, Mark C.
; TITLE OF INVENTION: Methods for Diagnosing and Treating Heart Disease
; FILE REFERENCE: 00786/381002
; CURRENT APPLICATION NUMBER: US/09/759,508B
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/175,787
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 26926
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-508B-2

Query Match          7.6%  Score 75;  DB 10;  Length 26926;
Best Local Similarity 22.3%;  Pred. No. 3.9e+03;
Matches 39;  Conservative 27;  Mismatches 67;  Indels 42;  Gaps 8;

QY 54 LYALARTTS-----PKVQRADSDVI-----FNSNFGERNVVTEGDKKVLGGCAPL 101
Db 21531 VYTVKAKNSGSAKAEIKVKVQDTPGKWGPRTFTNTEKMTJWDAFLN---DGCAPL 21587

QY 102 TRTNKLRIFGRTFTAYVDFCI-----YKHKLQLNAAAEELGIPAECSVL 148
Db 21588 THVIEKRETSRLJAWALIEDKCEAQSYTAIKLINGNEYQFRVSANVKFG-VGRPLSDSFV 21646

QY 149 AADFLGTCPKLSELQSRKQKFAWMAL-----KTEGGVWNTPVSNLRQ---LGRRE 196
Db 21647 VAGIQYTFDAPGCPPEPSNTGNSITLTWARPESDGG-----SEIQQVILERE 21695

RESULT 9
US-10-156-761-10849
; Sequence 10849, Application US/10-156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
```

```
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10849
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10849

Query Match          7.5%  Score 74.5;  DB 15;  Length 292;
Best Local Similarity 28.3%;  Pred. No. 5.5;
Matches 39;  Conservative 12;  Mismatches 48;  Indels 33;  Gaps 8;

QY 31 EVENMLVQKAPNEG-EVVFGLLYALAAATTSKVKQRADSDVIFSNFSGERNVVTEGC 90
Db 26 DMHTYLLGGWDERGAEAVYGFLEAAGQAGR---CRAAGDAAVVGC-----LLVDEGC 77

QY 91 -----LKKVLGGCAPL-----TRTNKLRIFGRT-----FTAYVDFCI 124
Db 78 GAGOFARYEAVLRKVAD-CAPVLLVPLSERF--DYGALGRVDALLVCGSLTFAYQDALA 134

QY 125 AYKHKLQLNAAAEELGIP 142
Db 135 EVLGRLPV--LAERGIP 150

RESULT 10
US-10-205-823-377
; Sequence 377, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 377
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-205-823-377

Query Match
Best Local Similarity 22.5%; Score 74; DB 15; Length 368;
Matches 50; Conservative 36; Mismatches 70; Indels 66; Gaps 12;

QY 24 DKLLSAEVEKMLVQKAPNEGIEV---FGLLLYALAARTTSPKVRADSVIFSNF 79
Db 130 DGRLEVDIDEVVDSPQNIKILHSKQFGNIL-----ILSGDVNLAESDLAYTRAI 183
QY 80 -----GERNVVTEG-----LK-----KVLGGCAPLRTFT--NKL 108
Db 184 MSGGKEDYTGKVLILGGGGGILCEIVLKPKRMVTWVIDQWIDGCKKXKTKCGDVL 243
QY 109 RTFGRTTTEAYVDFCI-----AYKHKLPLQNAAAELGIPAEDSYLAADFLGTCP 157
Db 244 DNLKGDCYQVLIEDCIPLVKRYAKEGREFDYVINDLTAVPISTSPEDS--TWEFLRLIL 301
QY 158 KLS--ELQQRKMFASMAKTEGGVNV-TPVSNL--RQLGR 194
Db 302 DLSMKVLKQDGKYP-----TQNCVNLTEALSLEYEQLR 336

RESULT 11
US-09-925-301-1385
; Sequence 1385, Application: US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05982
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/24,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent In Ver. 2.3
; SEQ ID NO 1385
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-301-1385

Query Match
Best Local Similarity 22.5%; Score 74; DB 9; Length 399;
Matches 50; Conservative 36; Mismatches 70; Indels 66; Gaps 12;

QY 24 DKLLSAEVEKMLVQKAPNEGIEV---FGLLLYALAARTTSPKVRADSVIFSNF 79
Db 161 DGRLEVDIDEVVDSPQNIKILHSKQFGNIL-----ILSGDVNLAESDLAYTRAI 214
QY 80 -----GERNVVTEG-----LK-----KVLGGCAPLRTFT--NKL 108
Db 215 MSGGKEDYTGKVLILGGGGGILCEIVLKPKRMVTWVIDQWIDGCKKXKTKCGDVL 274
QY 109 RTFGRTTTEAYVDFCI-----AYKHKLPLQNAAAELGIPAEDSYLAADFLGTCP 157
Db 275 DNLKGDCYQVLIEDCIPLVKRYAKEGREFDYVINDLTAVPISTSPEDS--TWEFLRLIL 332
QY 158 KLS--ELQQRKMFASMAKTEGGVNV-TPVSNL--RQLGR 194
Db 333 DLSMKVLKQDGKYP-----TQNCVNLTEALSLEYEQLR 367

RESULT 12
US-10-156-761-12518
; Sequence 12518, Application US/10156761
; Publication No. US20030119318A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATCSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HIROSHI, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12518
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
; US-10-156-761-12518

Query Match
Best Local Similarity 21.4%; Score 73.5; DB 15; Length 451;
Matches 42; Conservative 33; Mismatches 60; Indels 61; Gaps 10;

QY 1C SNLVTIDASSLNGVDKLLLSA-EVERKMLVQKAPNEGIEVFGLLLYALAARTTSPKVR 68
Db 138 AKLYIND-YNVEGVNAKSTALYNLVLKSLKORGVPIDGVGLQAHVLVGGVPS-TLQONIOR 255
QY 69 -ADSDVIFNSFGERNVVTET-----GSLKKVLGGCAPLRTFTNKL 109
Db 256 FADLGV-----DVAITELDIRMQLPATDACLTOAADYKAVLACVAVTRCVG-VT 305
QY 110 TFGRTTTEAYVDFCIAYKHKLPLQNAAAELGIPAEDSYLAADFLGTCPKLSLQSRMF 169
Db 306 VMGFTDSDWI-----PDVFSGYGAATPDENY-----APK----- 336
QY 170 ASMYALKTEGGVNVNTP 185
Db 337 PAYHAATAGGTSTP 352

RESULT 13
US-09-844-281-1
; Sequence 1, Application US/09844281
; Patent No. US20020082386A1
; GENERAL INFORMATION:
; APPLICANT: Mangold, Beverly L.
; APPLICANT: Aldrich, Jennifer L.
; APPLICANT: O'Brien, Thomas
; TITLE OF INVENTION: Anthrax Specific Antibodies
; FILE REFERENCE: 39632.0033
; CURRENT APPLICATION NUMBER: US/09/844,281
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,505
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 1
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; US-09-844-281-1

Query Match
Best Local Similarity 7.4%; Score 73.5; DB 9; Length 833;
Matches 53; Conservative 29; Mismatches 90; Indels 43; Gaps 10;

QY 8 NLSNLVITDASSLNGVDKLLLSAEVEKMLVQK-----APNEGIEVFGLLLYALAA--R 60
```

```

403 NSQNLVJGEKASLN---KLVATIAGEKVVDPGSGISIKSSNHGLIISVNNVYTAEGA 459
Ddb

61 TTSPKVRQADSDVIFPSNSFGERNVVTGEGD---LKKVLGDGCAPLTRFTNKLRTFGRTFE 117
Qy

460 TLTIKVGQVTHKDVFEKVTTDSRKULSVKAMPDKLVQVQNKTLPVTFVT-----TD 509
Ddb

118 AYVD-----FCIAYKHKLQPLNAAAEIG---PABDSYLAADFGLTGPCLSELQCSRKF- 169
Qy

510 QYGDFFGANTAAIKVLPTKTGVAAEGLDVVTDSGSGITKTKICVTG--NDVGEGTVHFQ 567
Ddb

170 -----ASWYALKTEGCVVNTP---VSNLRQLGR 194
Qy

568 KNGATIGSLYRVNVTGKVAFKNFELNSKVGYGO 602
Db

```

```

RESULT 14
US-10-156-761-13439
; Sequence 13439, Application US/1015676-
; Publication No. US200319019A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASASHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13439
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13439

```

```

Query Match      7.4% Score 73; DB 15; Length 485;
Best Local Similarity 23.3%; Pred. No. 17;
Matches 54; Conservative 3; Mismatches 79; Indels 68; Gaps 11;

      QY      5 SDSNLSLVITDASSLVGVDKLLSAEVEKMLVQK-----GAPNEG-EVVFGLLLYALA 58
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      Ddb      87 SDTQSSDVMVTCGG-----QALKSS-VEKVAQKGVASAVGALSINVTKVDGSGFTQCKA 140

      QY      59 ARTTS-----PKYQ--RAPSDEVIFSFGERNVVVTEGDLKKVLGDC 98
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      Ddb      141 KSSTTQGGGGGGGGGGTGTGFQVQGGGAGFDV--NSYVAGVDVTHQDL----- 190

      QY      99 APLRFNFNKLRTFGRTTEAYVDFCIAYKHKLPLQNAAAE-----LG 140
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      Ddb      191 GPLA--TSKI-TKGKTTT-SQTNAKVAVLK-----SYAKENKVTGKTIKSGTKYVTVIG 243

      QY      :41 IPABD-SYLAADFLGTCPKLUSELQQSRKMFASFYALKTTEGGVVNTFVSNLRQ 191
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      Ddb      244 IATPDSSESTTDVYLPKQAOATGDSGNKVTIIVKATDSKOIDRVKTIQK 295
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15  
UG-10-032-585-7518  
; Sequence 7518, Application US/10032585  
; Publication No. US20030180953A1  
; GENERAL INFORMATION:  
; APPLICANT: Terry, Roemer D.  
; APPLICANT: Bo, Jiang  
; APPLICANT: Charles, Boone  
; APPLICANT: Howard, Bussey  
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery

```

? FILE REFERENCE: 10192-005-999
? CURRENT APPLICATION NUMBER: US/10/032.585
? CURRENT FILING DATE: 2001-12-20
? NUMBER OF SEQ ID NOS: 8000
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 7518
? LENGTH: 1987
? TYPE: PRT
? ORGANISM: Candida albicans
US-10-032-585-7518

Query Match          7.4%; Score 73; DB 12; Length :987;
Best Local Similarity 21.4%; Pred. No. 1.4e+02;
Matches 46; Conservative 29; Mismatches 88; Indels 52; Gaps 7;

QY      21 NGVDKLLGAEVKMLVQGAP-----NEGII-----EYVFGLLLVALAARTTSPKYCRA 69
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      723 SGAWVOLSPQDILQGRAGRPYCRNGEGIIITSGDEVQVYLAILNQQLFISEQLIHKL 782
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||:

QY      70 DSVDFISNFGESNVVTGGDLKKVLDCGAPT-----RFTNKL 158
       ::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      783 VDMNI-----NAEIVAGSIITIEEG-EWLTYCYFFVRNLQSPALYGV EATYCF--NDP 833
       ::||::||::||::||::||::||::||::||::||::||::||::||::||:

QY      109 RTGRTFTAYVDPCFAYKHKLPLQAKA-----ABLGTPAESDYLAADFLGTCPKLSEL 162
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      834 TLYNRADIIVAFCLLHENKLVYNAALGSVASTELGXIAASHFYINFEINDIYGXMKLP 893
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||:

QY      163 QQSRKMFASNYALKTEGGVVNTPVSNLRQIGREV 197
       :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db      894 WHSETDLSVFS--NSGEFKYVPV---RQEERLEI 923
       :|::||::||::||::||::||::||::||::||::||::||::||::||:

Search completed: November 7, 2003, 14:55:17
Job time : 30 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2003, 14:44:53 ; Search time 41 seconds  
(without alignments)  
765.533 Million cell updates/sec

Title: US-09-613-486-15  
Perfect score: 991  
Sequence: 1 MELMSDSNLSNLVTDASSL.....GGVNTFVSNRLQGRREVX 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_15Jun03.\*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	198	20	AAW73482 Grapevine leafroll
2	286	28.9	204	16	AAR72682 Sugar beet yellows
3	89.5	9.0	499	23	ABG93273 C. albicans BAX-as
4	86.5	8.7	328	23	ABG91963 Herbicidally activ
5	86	8.7	295	21	AAG20735 Arabidopsis thalia
6	86	8.7	295	21	AAG39167 Arabidopsis thalia
7	86	8.7	360	21	AAG20734 Arabidopsis thalia
8	86	8.7	360	21	AAG39166 Arabidopsis thalia
9	86	8.7	360	23	ABG91962 Herbicidally activ

10	86	8.7	366	21	AAG20733 Arabidopsis thalia
11	86	8.7	366	21	AAG39165 Arabidopsis thalia
12	86	8.7	447	22	AAB96072 Putative P. abyssii
13	85	8.6	599	18	AAW17788 Prage abortive inf
14	84.5	8.5	269	22	AAV97690 GLAV-5 coat prote
15	81.5	8.2	402	21	AAAB4800 Erysipelothrix rhu
16	81.5	8.2	402	23	ABG07787 E. rhusiopathiae pr
17	81.5	8.2	402	23	ABG08869 Erysipelothrix rhu
18	81.5	8.2	626	21	AAV95782 Erysipelothrix rhu
19	80	8.1	555	23	ABBS4252 Lactococcus lactis
20	80	8.1	711	23	ABP29995 Streptococcus poly
21	80	8.1	729	23	ABP25672 Rat hemidesmosome
22	79.5	8.0	770	18	AAW26584 Propionibacterium
23	79	8.0	263	22	AAU49697 Enterococcus faeca
24	79	8.0	451	22	AAU35128 Grapevine leafroll
25	79	8.0	477	18	AAW21643 Drosophila melanog
26	79	8.0	782	22	ABBS7835 Lactococcus lactis
27	78.5	7.9	208	23	ABBS4384 Candida albicans e
28	77	7.8	696	23	ABP73971 Novel human secret
29	76.5	7.7	200	22	AAU30591 Grapevine leafroll
30	76	7.7	223	20	AAW73481 NADH oxidase, Amp
31	75.5	7.6	509	15	AAW47583 E. rhusiopathiae s
32	75.5	7.6	606	20	AAV43219 Herbicidally activ
33	75	7.6	358	23	ABG91943 Streptococcus poly
34	75	7.6	404	23	ABP25433 Synechococcus CWPB
35	75	7.6	883	22	AAW51651 Human titin (conne
36	75	7.6	26926	22	AAU05396 Human RGS11 protei
37	75	7.6	31267	24	ABG74786 Beta-TRCP.N/SK22.C
38	74.5	7.5	449	22	ABW48289 Rat laminin 5 poly
39	74.5	7.5	1694	21	AAAB48461 Rat laminin 5 poly
40	74.5	7.5	1725	21	AAAB48460 Herbicidally activ
41	74	7.5	359	23	ABG91946 Human cancer assoc
42	74	7.5	399	21	AAAB43940 Arabidopsis thalia
43	73.5	7.4	332	21	AAW10781 Arabidopsis thalia
44	73.5	7.4	355	21	AAW10780 Arabidopsis thalia
45	73.5	7.4	367	21	AAW10779 Arabidopsis thalia

ALIGNMENTS

RESULT 1  
AAW73482  
ID AAW73482 standard; Protein; 198 AA.  
XX  
AC AAW73482;  
XX  
CT 29-MAR-1999 (first entry)  
XX  
DE Grapevine leafroll virus type 2 coat protein.  
DE  
XX  
KW GLAV-2; closterovirus; grape; tobacco; transgenic plant;  
KW disease resistance; virus resistance; beet yellows virus;  
KW tristeza virus; coat protein.  
XX  
OS Grapevine leafroll virus type 2.  
XX  
PN WO9853055-A1.  
XX  
PD 26-NOV-1998.  
XX  
PF 20-MAY-1998; 98WO-US10313.  
XX  
PR 20-MAY-1997; 97US-0047194.  
XX  
PA (CCRR ) CORNELL RES FOUND INC.  
XX  
PI Gonsalves D, Ling K, Zhu H;  
XX  
DR WP2; 1999-045307/04.  
XX  
N-PSDB; AAV08870.  
XX  
PT Grapevine leafroll virus (type 2) proteins and polypeptides - and

PT encoding DNA, useful e.g. to impart grapevine leafroll resistance to  
 PT grape and tobacco plants and detect grapevine leafroll virus  
 XX  
 PS Claim 12; Page 44-45; 151pp; English.

XX This is the amino acid sequence of a 22 kDa coat protein that is  
 CC encoded by open reading frame ORF6 (see AAV08870) of grapevine  
 CC leafroll virus type 2 (GIRAV-2) RNA (see AAV08874). The GIRAV-2  
 CC genome includes 9 open reading frames (see AAV3864-72) for a  
 CC polyprotein, an RNA-dependent RNA polymerase, heat shock proteins,  
 CC coat proteins and proteins of unknown function (see AAV73476-84).  
 CC These can be used to produce antibodies useful for detecting  
 CC GIRAV in samples e.g. by ELISA (claimed). The nucleic acid  
 CC molecules can be used to produce probes and primers for such  
 CC detection, and to transform host cells (especially Agrobacterium  
 CC vitis, Agrobacterium tumefaciens, grape, citrus, beet or tobacco  
 CC cells) and produce transgenic plants (claimed). They can be used  
 CC to impart GIRAV-2 resistance to Vitis scion or rootstock cultivars  
 CC or Nicotiana cultivars (claimed). Because extensive similarity  
 CC exists between hsp70-related sequence regions of GIRAV-2 and other  
 CC closteroviruses, the DNA may also be used to impart beet yellows  
 CC virus resistance to beet cultivars or tristera virus resistance to  
 CC citrus scion cultivar/rootstock cultivars (claimed).

XX  
 SQ Sequence 198 AA;  
 Query Match 100.0%; Score 991; DB 20; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-100;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY : MELMSDNLNLVITDASSLNGVDKLLSAEVEKMLVQKGNPEGIEVFGGLLYALAAAR 60  
 DB 1 MELMSDNLNLVITDASSLNGVDKLLSAEVEKMLVQKGNPEGIEVFGGLLYALAAAR 60  
 QY 61 TTSFKVQRADSDVIFNSFGERNVNVTEGDLKKVLDGCAPLTRFTNKLRTFGRTTEAYV 120  
 DB 61 TTSFKVQRADSDVIFNSFGERNVNVTEGDLKKVLDGCAPLTRFTNKLRTFGRTTEAYV 120  
 QY 121 DFCIAYKHKLPLQNLAAELGIPAEDSYLAADFCTGCPKLSLQOSKMPFASMYALKTEGG 180  
 DB 121 DFCIAYKHKLPLQNLAAELGIPAEDSYLAADFCTGCPKLSLQOSKMPFASMYALKTEGG 180  
 QY 181 VNTPTVSNLRQLGRREV 198  
 DB 181 VNTPTVSNLRQLGRREV 198

RESULT 2  
 AAR72682  
 ID AAR72682 standard; Protein: 204 AA.  
 XX  
 AC AAR72682;  
 DT 25-MAR-2003 (updated)  
 DT 02-NOV-1995 (first entry)  
 XX  
 XX Sugar beet yellows virus capsid protein.  
 DE  
 XX  
 XX Sugar beet yellows virus capsid protein; transgenic plant.  
 KW  
 OS Sugar beet yellows virus.  
 XX  
 XX RUI2017820-C1.  
 XX  
 XX 15-AUG-1994.  
 XX  
 XX 27-JUN-1991; 9ISU-4950054.  
 XX  
 XX 27-JUN-1991; 9ISU-4950054.  
 XX  
 XX (BIOT=) BIOYCHN INST CO LTD.  
 PA (IMMU=) IMMUNOBIOTECHN INST.  
 XX  
 XX

P: Arganovskii AA, Boiko VP, Karasev AV;  
 XX WPI; 1995-113715/15.  
 DR N-PSDB; AAQ87853.  
 XX  
 XX Sugar beet yellows virus cDNA fragment encoding capsid protein -  
 PT useful for production of virus-resistant transgenic plants  
 PT  
 XX Claim 1; Column 7-10; 5pp; Russian.  
 XX  
 CC The amino acid sequence of the novel sugar beet yellows virus (SBYV)  
 CC capsid protein. The protein has mol. wt. 22.2 kD. The corresponding  
 CC gene was obtained from reverse transcribed RNA isolated from purified  
 CC SBYV. The fragment is useful for the production of virus resistant  
 CC transgenic plants by genetic engineering methods.  
 CC (Updated on 25-MAR-2003 to correct EN field.)  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 204 AA;  
 Query Match 28.9%; Score 286; DB 16; Length 204;  
 Best Local Similarity 37.6%; Pred. No. 6.2e-23;  
 Matches 71; Conservative 29; Mismatches 87; Indels 2; Gaps 2;  
 QY 7 SNLSNLVITDASSLNGVDKLLSAEVEKMLVQKGNPEGIEVFGGLLYALAAATTSFKV 66  
 DB 11 ATFNVSIAQOTCLHGDCDCKLRKNFECLKLVGPEDNLGIALGLCLYSCATIGTSNKV 70  
 QY 67 QRASDVIFNSF-GERNVNVTEGDLKKVLDGCAPLTRFTNKLRTFGRTTEAYVDFCIA 125  
 DB 71 NVQPTSF-PIKASFGGKELVLTGELNSFLGSKLLEKPKKJRCFCRTFKDYISJRKE 130  
 QY 126 YKHKLPLQNLAAELGIPAEDSYLAADFCTGCPKLSLQOSKMPFASMYALKTEGGVNT? 185  
 DB 131 YGKLPPIARANRGLPAECHYLAADFISTSTELTDLQOSKLLARENATHTEFS-SESP 189  
 QY 186 VSNLRQLGR 194  
 DB 190 VTSJQLGR 198  
 RESULT 3  
 ABG93273  
 ID ABG93273 standard; Protein: 499 AA.  
 XX  
 AC ABG93273;  
 XX  
 DT 21-NOV-2002 (first entry)  
 XX  
 DE C. albicans BAX-associated protein fragment SEQ ID 504.  
 XX  
 KW Bax: Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;  
 KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;  
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;  
 KW neurodegeneration; cell death.  
 XX  
 CS Candida albicans.  
 XX  
 XX WO200264766-A2.  
 XX  
 PD 22-AUG-2002.  
 XX  
 PF 21-DEC-2001; 2001WO-EP15398.  
 XX  
 XX 22-DEC-2000; 2000EP-C870318.  
 ER 04-JAN-2001; 2001EP-C870002.  
 XX 09-JAN-2001; 2001EP-C870003.  
 XX  
 PA (CANC) JANSSEN PHARM NV.  
 XX  
 XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RC;  
 P: WPI; 2002-667002/71.  
 DR



PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 08-MAY-1999; 99US-0132048.  
PR 14-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137526.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0142977.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 25-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 36-OCT-1999; 99US-0157865.

PR	07-OCT-1999;	99US-0158029.	PF	25-FEB-2003;	200EP-0301439.
PR	08-OCT-1999;	99US-0158232.	XX		
PR	12-OCT-1999;	99US-0158369.	PR	25-FEB-1999;	99US-0121625.
PR	13-OCT-1999;	99US-0159293.	PR	05-MAR-1999;	99US-0123180.
PR	13-OCT-1999;	99US-0159294.	PR	09-MAR-1999;	99US-0123548.
PR	13-OCT-1999;	99US-0159295.	PR	23-MAR-1999;	99US-0125786.
PR	14-OCT-1999;	99US-0159329.	PR	25-MAR-1999;	99US-0126264.
PR	14-OCT-1999;	99US-0159330.	PR	29-MAR-1999;	99US-0126785.
PR	14-OCT-1999;	99US-0159331.	PR	01-APR-1999;	99US-0127462.
PR	14-OCT-1999;	99US-0159637.	PR	06-APR-1999;	99US-0128234.
PR	18-OCT-1999;	99US-0159638.	PR	08-APR-1999;	99US-0128714.
PR	21-OCT-1999;	99US-0159584.	PR	16-APR-1999;	99US-0129845.
PR	21-OCT-1999;	99US-016074.	PR	19-APR-1999;	99US-0130077.
PR	21-OCT-1999;	99US-0160767.	PR	21-APR-1999;	99US-0130449.
PR	21-OCT-1999;	99US-0160768.	PR	23-APR-1999;	99US-0130510.
PR	21-OCT-1999;	99US-0160770.	PR	23-APR-1999;	99US-0130891.
PR	21-OCT-1999;	99US-0160814.	PR	28-APR-1999;	99US-0131449.
PR	21-OCT-1999;	99US-0160815.	PR	30-APR-1999;	99US-0132048.
PR	22-OCT-1999;	99US-0160980.	PR	30-APR-1999;	99US-0132407.
PR	22-OCT-1999;	99US-0160981.	PR	04-MAY-1999;	99US-0132484.
PR	22-OCT-1999;	99US-0160989.	PR	05-MAY-1999;	99US-0132485.
PR	25-OCT-1999;	99US-0161404.	PR	06-MAY-1999;	99US-0132486.
PR	25-OCT-1999;	99US-0161405.	PR	06-MAY-1999;	99US-0132487.
PR	25-OCT-1999;	99US-0161406.	PR	07-MAY-1999;	99US-0132863.
PR	26-OCT-1999;	99US-0161359.	PR	11-MAY-1999;	99US-0134256.
PR	26-OCT-1999;	99US-0161360.	PR	14-MAY-1999;	99US-0134218.
PR	26-OCT-1999;	99US-0161361.	PR	14-MAY-1999;	99US-0134219.
PR	28-OCT-1999;	99US-0161920.	PR	14-MAY-1999;	99US-0134221.
PR	28-OCT-1999;	99US-0161992.	PR	14-MAY-1999;	99US-0134370.
PR	28-OCT-1999;	99US-0161993.	PR	18-MAY-1999;	99US-0134768.
PR	29-OCT-1999;	99US-0162142.	PR	19-MAY-1999;	99US-0134941.
			PR	20-MAY-1999;	99US-0135124.
			PR	21-MAY-1999;	99US-0135353.
			PR	24-MAY-1999;	99US-0135629.
			PR	25-MAY-1999;	99US-0136021.
			PR	27-MAY-1999;	99US-0136392.
			PR	28-MAY-1999;	99US-0136782.
			PR	31-JUN-1999;	99US-0137222.
			PR	03-JUN-1999;	99US-0137528.
			PR	04-JUN-1999;	99US-0137502.
			PR	07-JUN-1999;	99US-0137724.
			PR	08-JUN-1999;	99US-0138094.
			PR	10-JUN-1999;	99US-0138540.
			PR	14-JUN-1999;	99US-0138847.
			PR	16-JUN-1999;	99US-0139119.
			PR	16-JUN-1999;	99US-0139452.
			PR	16-JUN-1999;	99US-0139453.
			PR	17-JUN-1999;	99US-0139492.
			PR	18-JUN-1999;	99US-0139454.
			PR	18-JUN-1999;	99US-0139455.
			PR	18-JUN-1999;	99US-0139456.
			PR	18-JUN-1999;	99US-0139457.
			PR	18-JUN-1999;	99US-0139458.
			PR	18-JUN-1999;	99US-0139459.
			PR	18-JUN-1999;	99US-0139460.
			PR	18-JUN-1999;	99US-0139461.
			PR	18-JUN-1999;	99US-0139462.
			PR	18-JUN-1999;	99US-0139463.
			PR	18-JUN-1999;	99US-0139750.
			PR	18-JUN-1999;	99US-0139763.
			PR	21-JUN-1999;	99US-0139817.
			PR	22-JUN-1999;	99US-0139899.
			PR	23-JUN-1999;	99US-0140353.
			PR	23-JUN-1999;	99US-0140354.
			PR	24-JUN-1999;	99US-0140695.
			PR	28-JUN-1999;	99US-0140823.
			PR	28-JUN-1999;	99US-0140991.
			PR	30-JUN-1999;	99US-0141287.
			PR	01-JUL-1999;	99US-0141842.
			PR	01-JUL-1999;	99US-0142154.
			PR	02-JUL-1999;	99US-0142055.
			PR	26-JUL-1999;	99US-0142390.
			PR	08-JUL-1999;	99US-0142803.

Query Match 8.7%; Score 86; DB 21; Length 295;

Best Local Similarity 26.5%; Pred No. 0.94;

Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 3;

QY	7	SNLSNLVITDASSLNGV--DKKLSAEVEKMLVQKGAENEGI----EYVFGLLIYALAAAR	60
DB	72	SEQPNFKSYIARLKGIVGDKAMEIINNAFVVSAGPDLKLYEIPSRRLVYP--ISG	13:
QY	61	TTSRKVQADSDVIFSNFSGRNWVVTGDLKKVLDGCAELTRTNKLRITGRITFTAYV	120
DB	132	YQDFILKRLNFVRELYSLGVNRNLV--GGLPFM--GCLPI-HMTAKPRNIFRCLEHHN	186
QY	121	DFCIAKHKLPCLNAAELGLIPAEBSYLAADFGLTCTPKLSELQSRKMFASVYALK-TEG	179
DB	187	KDSVLYNEKQLNLLPQIEASLPG-SKFLYADVYN--PYMEMIQNPSK-----YGFKEIKR	238
QY	180	GVVNT	184
DB	239	GCCGT	243

RESULT 6

AG39167  
ID AG39167 standard; Protein; 295 AA.

XX AC  
AC AAG39167;

XX DT  
DT 18-OCT-2000 (first entry)

XX DE  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48421.

XX KW  
KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW  
KW termination sequence.

XX OS  
OS Arabidopsis thaliana.

XX PN  
PN EP1033405-A2.

XX PD  
PD 06-SEP-2000.

XX



```
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147322.
PR 05-AUG-1999; 99US-0147262.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151365.
PR 27-AUG-1999; 99US-0151366.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 8.7%; Score 86; DB 21; Length 295;
Best Local Similarity 26.5%; Pred. No. 3.94;
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLVITDASSLNGV--DKKLLSAEVEKMLVQKGNPNEGJ----EVVFGLLLYALAAAR 60
DB 72 SEQPNKFKSYIARLKGIVGCKKAMEIINNAFVVVSAGNDFILNYVEIPSRRLVYPPISG 131
QY 61 TTSPKVRADSDVIFNSFGERNVVVTEGELKKVLDGCAPLRTFNKLRFTGRTTTEAYV 120
DB 132 YQDFTLKRLNFVRELYSLGVRNLV--GGLPWK--GCLPI--HMTAKFRNIFRCLEHHN 186
QY 121 DFCIAYKHLQOLNNAEELGTPAEDSYLAADPLGTCPKLSELQOSRXXFASMYALK-TEG 179
DB 187 KDSVLNFKLONLQFQIEASLPG-SKFLYADVYN--PMMEMIQNPSK-----YGFKEKTR 238
QY 180 GVVNT 184
DB 239 GCCGT 243

RESULT 7
AAG20734
-D AAG20734 standard; Protein; 360 AA.
XX
XX AAG20734;
AC
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23036.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
```

[illegible]



```
PR 23-JUN-1999; 99JUS-0140354;
PR 24-JUN-1999; 99JUS-0140695;
PR 28-JUN-1999; 99JUS-0140823;
PR 29-JUN-1999; 99JUS-0140991;
PR 30-JUN-1999; 99JUS-0141287;
PR 01-JUL-1999; 99JUS-0141842;
PR 01-JUL-1999; 99JUS-0142154;
PR 02-JUL-1999; 99JUS-0142559;
PR 06-JUL-1999; 99JUS-0142390;
PR 08-JUL-1999; 99JUS-0142803;
PR 09-JUL-1999; 99JUS-0142920;
PR 12-JUL-1999; 99JUS-0142977;
PR 13-JUL-1999; 99JUS-0143542;
PR 14-JUL-1999; 99JUS-0143624;
PR 15-JUL-1999; 99JUS-0144005;
PR 16-JUL-1999; 99JUS-0144088;
PR 16-JUL-1999; 99JUS-0144086;
PR 19-JUL-1999; 99JUS-0144325;
PR 19-JUL-1999; 99JUS-0144331;
PR 19-JUL-1999; 99JUS-0144332;
PR 19-JUL-1999; 99JUS-0144333;
PR 19-JUL-1999; 99JUS-0144334;
PR 19-JUL-1999; 99JUS-0144335;
PR 20-JUL-1999; 99JUS-0144352;
PR 20-JUL-1999; 99JUS-0144632;
PR 20-JUL-1999; 99JUS-0144884;
PR 21-JUL-1999; 99JUS-0144814;
PR 21-JUL-1999; 99JUS-0145086;
PR 21-JUL-1999; 99JUS-0145088;
PR 22-JUL-1999; 99JUS-0145085;
PR 22-JUL-1999; 99JUS-0145087;
PR 22-JUL-1999; 99JUS-0145089;
PR 22-JUL-1999; 99JUS-0145192;
PR 23-JUL-1999; 99JUS-0145145;
PR 23-JUL-1999; 99JUS-0145218;
PR 23-JUL-1999; 99JUS-0145224;
PR 26-JUL-1999; 99JUS-0145276;
PR 27-JUL-1999; 99JUS-0145913;
PR 27-JUL-1999; 99JUS-0145918;
PR 27-JUL-1999; 99JUS-0145919;
PR 28-JUL-1999; 99JUS-0145951;
PR 02-AUG-1999; 99JUS-0146386;
PR 02-AUG-1999; 99JUS-0146388;
PR 02-AUG-1999; 99JUS-0146389;
PR 03-AUG-1999; 99JUS-0147038;
PR 04-AUG-1999; 99JUS-0147204;
PR 04-AUG-1999; 99JUS-0147302;
PR 05-AUG-1999; 99JUS-0147192;
PR 06-AUG-1999; 99JUS-0147260;
PR 06-AUG-1999; 99JUS-0147303;
PR 09-AUG-1999; 99JUS-0147416;
PR 09-AUG-1999; 99JUS-0147493;
PR 10-AUG-1999; 99JUS-0147935;
PR 11-AUG-1999; 99JUS-0148319;
PR 12-AUG-1999; 99JUS-0148341;
PR 13-AUG-1999; 99JUS-0148565;
PR 13-AUG-1999; 99JUS-0148684;
PR 16-AUG-1999; 99JUS-0149368;
PR 17-AUG-1999; 99JUS-0149175;
PR 18-AUG-1999; 99JUS-0149426;
PR 20-AUG-1999; 99JUS-0149722;
PR 20-AUG-1999; 99JUS-0149723;
PR 23-AUG-1999; 99JUS-0149929;
PR 23-AUG-1999; 99JUS-0149902;
PR 23-AUG-1999; 99JUS-0149220;
PR 25-AUG-1999; 99JUS-0150566;
PR 26-AUG-1999; 99JUS-0150884;
PR 27-AUG-1999; 99JUS-0151065;
PR 27-AUG-1999; 99JUS-0151066;
PR 27-AUG-1999; 99JUS-0151080;
PR 30-AUG-1999; 99JUS-0151303;
PR 31-AUG-1999; 99JUS-0151438;

PR 01-SEP-1999; 99JUS-0151930;
PR 07-SEP-1999; 99JUS-0152263;
PR 13-SEP-1999; 99JUS-0153070;
PR 13-SEP-1999; 99JUS-0153758;
PR 15-SEP-1999; 99JUS-0154018;
PR 16-SEP-1999; 99JUS-0154039;
PR 20-SEP-1999; 99JUS-0154779;
PR 22-SEP-1999; 99JUS-0155139;
PR 23-SEP-1999; 99JUS-0155486;
PR 24-SEP-1999; 99JUS-0155659;
PR 28-SEP-1999; 99JUS-0156458;
PR 29-SEP-1999; 99JUS-0156596;
PR 04-OCT-1999; 99JUS-0157117;
PR 05-OCT-1999; 99JUS-0157753;
PR 06-OCT-1999; 99JUS-0157865;
PR 07-OCT-1999; 99JUS-0158029;
PR 08-OCT-1999; 99JUS-0158232;
PR 12-OCT-1999; 99JUS-0158369;
PR 13-OCT-1999; 99JUS-0159293;
PR 13-OCT-1999; 99JUS-0159294;
PR 13-OCT-1999; 99JUS-0159295;
PR 14-OCT-1999; 99JUS-0159329;
PR 14-OCT-1999; 99JUS-0159330;
PR 14-OCT-1999; 99JUS-0159331;
PR 14-OCT-1999; 99JUS-0159637;
PR 14-OCT-1999; 99JUS-0159638;
PR 18-OCT-1999; 99JUS-0159584;
PR 21-OCT-1999; 99JUS-0160741;
PR 21-OCT-1999; 99JUS-0160767;
PR 21-OCT-1999; 99JUS-0160768;
PR 21-OCT-1999; 99JUS-0160770;
PR 21-OCT-1999; 99JUS-0160814;
PR 21-OCT-1999; 99JUS-0160815;
PR 22-OCT-1999; 99JUS-0160980;
PR 22-OCT-1999; 99JUS-0160981;
PR 22-OCT-1999; 99JUS-0160989;
PR 25-OCT-1999; 99JUS-0161404;
PR 25-OCT-1999; 99JUS-0161405;
PR 25-OCT-1999; 99JUS-0161406;
PR 26-OCT-1999; 99JUS-0161359;
PR 26-OCT-1999; 99JUS-0161360;
PR 26-OCT-1999; 99JUS-0161361;
PR 28-OCT-1999; 99JUS-0161920;
PR 28-OCT-1999; 99JUS-0161992;
PR 28-OCT-1999; 99JUS-0161993;
PR 23-OCT-1999; 99JUS-0162142;

Query Match      8.7%; Score 86; DB 21; Length 360;
Best Local Similarity 26.5%; Pred. No. 1.2;
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY      7 SNLSNLVITDASSLNGV--DKKLSAEVEKMLVQKAPNEG1---EVVFGLLLYALAAR 60
Db      137 SEQPNFKSVIARLKGIVGKKAMEIINNAFVVVSAGPNDFLNYYEIPSRRLVYFFISG 196

QY      61 TTSPKVRADSDVIFNSFGERNVVVTEGDKKVLDDGAPJTRFTNKJLRTFORTFTTEAYV 120
Db      197 YQCFILKRLNFVRELYSLGRNVLV--GGLPPM--GCLPI-HM-AKFRNIFRCLEHHN 251

QY      121 DFCIAYKHLIPLQNAAEELGIPAEDSYLAADFLGTCFKLSE:QQSKKMFASVYALK-TEG 179
Db      252 KDSVLYNEKLQNLPLQIEASLP3-SKFLYADVYN--PMMEMIONPSK-----YGFKEIKR 303

QY      180 GVYNT 184
Db      304 GCCGT 308

RESULT 9
ABB91962
ID ABB91962 standard; Protein; 360 AA.
XX ABB91962;
AC ABB91962;
```



PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0156659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159638.
PR	20-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0144814.	PR	22-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160767.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160915.
PR	22-JUL-1999;	99US-0145192.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160989.
PR	23-JUL-1999;	99US-0145224.	PR	25-OCT-1999;	99US-0161404.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145918.	PR	26-OCT-1999;	99US-0161359.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161360.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146386.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161992.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161983.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147204.			
PR	04-AUG-1999;	99US-0147302.			
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149366.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151066.			
PR	27-AUG-1999;	99US-0151080.			
PR	30-AUG-1999;	99US-0151303.			
PR	31-AUG-1999;	99US-0151438.			

Query Match

Best Local Similarity 8.7%; Score 86; DB 21; Length 366;

Mismatches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY	7	SNLSNLVITDASSLNGV--DKKLLSREVEKMLVQKAPNEG-----EVVFGULLYLAAR	60
DB	143	SEQPMFKSYIAKLGKIVGDKKAME--NNAFVVSAGPKDF--LNYYE--FSRRLEYFP--SG	202
QY	61	TSPPKQVQADSDVIFNSFGERNVVVTEGDKKVLDCAPLTFRTNKLRTFGRTF-EAYV	120
DB	203	YQDFILKRLNFVRELYSLGVRNLV--GGLPPX--GCLPI--HMTAKFRN--FAPCLEHNN	257
QY	121	DFCIAYKHLKPOLNAAAEIGIPAEDSYLAADFGLGTPKUSELQOSKRYFASMYAJK--TEG	179
DB	259	KDSVLVNEK--QNLPLQIEASLPQ--SKFLYADVNN--PMMEYIQNPCK-----YGFNETKR	309
QY	180	GVVNT 184	
DB	310	GCCGT 314	

RESULT 11

AAG39165

ID AAG39165 standard; Protein; 366 AA.

XX

AC AAG39165;



```
PR 27-AUG-1999; 99US-0131065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154775.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155466.
PR 24-SEP-1999; 99US-0156658.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-016074.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161362.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query March 8.7%; Score 86; DB 21; Length 366;
Best Local Similarity 26.5%; Pred. No. 1.3;
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLTVDASSNGV--DKKLSAEVEKMLVQKAPNEG1----EVVFGLLVALAAR 60
DB 143 SEQPNFYSYARLKGIVGDKAMEIINNAFVVSAGNDFILNYEIPSRLEYPPISG 202
QY 61 TSPKVRADSDVIFSNFSGERNVVVTEGDLKKVLDGCAPLRTFNKLRTEGRTFTAYV 120
DB 203 VQDFILKLENFVRELYSLGVNVLV--GGLPPI--GCLPI-HMTAKFNI-FRFLCEHHN 257
QY 121 DFCIAYKHLKLPOLNAAAE--GIPAEOSYIAADFLGTCPKLSLELQQRKMFASKYALK-TEG 179
DB 258 KDSVLYNEKLNQLPQIEASLPG-SKFLYADVYN--PMWEMIQNPSK-----YGFKETKR 309
QY 180 GVNT 184
DB 310 GCCGT 314
```

```
RESULT 12
AAB96072
ID AAB96072 standard; Protein; 447 AA.
XX
AC AAB96072;
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative P. abyssi beta-lactamase.
XX
KW Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS Pyrococcus abyssi.
XX
PR 2792651-A1.
XX
PD 27-OCT-2000.
XX
PF 21-APR-1999; 99PR-0005034.
XX
PR 21-APR-1999; 99PR-0005034.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (IFRE-) IPREMER INST FR RECH EXPL MER.
XX
PI Forterre P, Thierry JC, Prieur J, Dietrich J, Lecompte O;
PI Querellou J, Weissenbach J, Saurin W, Heilig R;
XX
XX WPI; 2001-126236/14.
XX
XX New nucleotide sequences isolated from Pyrococcus abyssi encode
PT proteins useful in industry -
XX
XX Claim 7; Pages 693-694; 1657pp; French.
XX
XX The present invention relates to the genomic sequence of Pyrococcus
XX abyssi (see AAF86431 and AAF41223-7) and P. abyssi proteins. P. abyssi is
XX a hyperthermophilic archaeon, which is isolated from deep-sea
XX hydrothermal vents. The present sequence is one such P. abyssi protein.
XX The proteins of the present invention have various potential industrial
XX uses, since the proteins are stable at very high temperatures, some up to
XX 110 degrees centigrade.
XX Note: This patent is in the same patent family as WO2000065062, which
XX contains additional sequences as shown in AAB99132-AAB99143,
XX AAH75903-AAH75920 and AAG66436.
XX
SQ Sequence 447 AA;
Query March 8.7%; Score 86; DB 22; Length 447;
Best Local Similarity 26.0%; Pred. No. 1.7;
Matches 40; Conservative 15; Mismatches 41; Indels 58; Gaps 8;

QY 26 KLSAEVSKLVOKGAPNEG-EVVFGLLYVALAARTSPKVRADSDVIFSNFSGERNV- 84
DB 5 KLSFIVEKVAERK-VPGISISII-----KGDVVYAKGFGYRNVE 44
QY 85 -----VVTEGDLKKVLDGCAPLRTFTN-KLRTFGRTFT 116
DB 45 ARLPSTPETIYGIGSTKSFALAIKLVLEGSL--LDD--PVEKFNKLRPFGEPT 100
QY 117 EAYVDFCIAYKHLKLPOLNAAAE--GIPAEOSYL 148
DB 101 ---VHLLTHSSGIPSLGYAEAFIDGMVSGDNWL 131
RESULT 13
AAW17788
ID AAW17788 standard; Protein; 599 AA.
XX
AC AAW17788;
XX
DT 19-AUG-1997 (first entry)
XX
```



```

DE Phage abortive infection protein AbiE.
XX
XX
XX AbiE; abortive phage infection protein; phage resistance;
KW PSRQ800; lactic acid bacterium.
XX
XX Lactococcus lactis subsp. lactis isolate W1.
OS WO3720917-A2.
XX
XX
XX WO3720917-A2.
XX
XX 12-JUN-1997.
XX
XX 20-NOV-1996; 96WO-IB01385.
XX
XX 01-DEC-1995; 95US-0565907.
XX
XX (UNIL ) QUEST INT BV.
XX
XX Holier BJ, Kondo JK, Moineau S, Vandenbergh PA;
PI Vedamuthu ER;
XX
XX WPI: 1997-319765/29.
XX
XX N-PSDB; AA768648.
XX
XX Isolated DNA encoding the AbiE protein of Lactococcus - for
PT protecting strains used in production of fermented dairy products
XX
XX Claim 48; Page 30-32; 49pp; Engl:sh.
XX
XX AbiE (AA017798) is a phage abortive infection protein that increases
CC resistance to phage. It is the expression product of an open
CC reading frame found in the 4.5 kb EcoRI fragment (AA168648) of
CC Lactococcus lactis (L.L.; subsp. lactis W1 plasmid PSRQ800. AbiE
CC acts at, or before, phage replication and has no homology with
CC known Abi proteins. It protects against phages of 936 and P335
CC types, also against c2 when present on a high copy number plasmid.
CC AbiE can be used to impart phage resistance to bacteria, esp. L1.
CC that do not have natural resistance. The recombinant bacteria can
CC be used in the prodn. of fermented dairy products.
XX
XX
XX Sequence 539 AA;
SQ
Query Match 8.6%; Score 85; DB 18; Length 599;
Best Local Similarity 24.3%; Pred. No. 3.3;
Matches 36; Conservative 26; Mismatches 46; Indels 40; Gaps 6;
QY 2 ELMSDSNL-----SNLIVTD-----ASSUNGVDK--KLSA 30
DB 261 EFLNEFNLCRENNLI:NDNKTKVDNPFVDSKSDIFSFFENITSNCKWKEISN 320
QY 31 EWEKMLVQKGAPNEG-IEVVFGLLYLA:AAATTSKVVORADSVIFNSFGERNVYVTEG 89
DB 321 FIDYCWNEEHGNGAKICFPVI-----TNTLKQKKVETKNI-DNIFSKRMVTFN 372
QY 90 DLKXVLDGCAPLRTFKLRTFGRTFE 117
DB 373 VFEKILDLSLKDSRLTNKFLTFEENINE 403
RESULT 14
AA097690
ID AA097690 standard; Protein; 269 AA.
XX
XX AA097690;
AC
XX
XX 08-MAY-2001 (first entry)
XX
XX GLRaV-5 coat protein.
XX
XX GLRaV-5; grapevine leafroll virus; GLRaV infection; GLRaV coat protein;
KW GLRaV HSP70 homologue protein; viral gene mapping;
XX plant disease resistance.
XX
XX Grapevine leafroll virus.
OS
XX
XX
XX WO250105957-A2.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US19708.
XX
XX 19-JUL-1999; 99US-0144453.
XX
XX (AGR1-) AGRITOPE INC.
XX
XX Good XC, Monis J;
XX
XX WPI: 2001-147339/15.
XX
XX N-PSDB; AA091259.
XX
XX Novel grapevine leafroll virus polynucleotide useful as diagnostic and
PT probe, for viral gene mapping and for induced plant disease resistance
XX
XX Claim 1; Fig 1; 60pp; English.
XX
XX This sequence represents a grapevine leafroll virus (GLRaV-5) protein
CC sequence of the invention. The DNA sequence can be used in an expression
CC construct. The construct is useful for providing resistance to GLRaV
CC infection in a recombinant plant cell by transforming the plant cell with
CC it, where transcription of the polynucleotide sequence interferes with a
CC normal viral function such as movement, encapsidation or replication of
CC viral RNA. The polynucleotide sequence is expressed as an antisense
CC sequence and encodes a GLRaV coat protein, preferably a defective GLRaV
CC coat protein or a GLRaV HSP70 homologue protein. The GLRaV-5 DNA is
CC useful for the synthesis of GLRaV, as diagnostics and probes, for viral
CC gene mapping and for induced plant disease resistance. It is also useful
CC to detect and quantitate expression of GLRaV in plant tissue prior to use
CC in vegetative propagation, by detecting the presence of GLRaV RNA.
XX
XX Sequence 269 AA;
SQ
Query Match 8.5%; Score 84.5; DB 22; Length 269;
Best Local Similarity 24.4%; Pred. No. 1.2;
Matches 32; Conservative 21; Mismatches 57; Indels 21; Gaps 4;
QY 61 TTSKVVORADSVIFNSFGERNVYVTEGDLKXVLDGCAPLRTFKLRTFGRTFEAYV 120
DB 127 STSPKVVSSNNRTI-TGYDQKEVTVAHDEIKTALDINSIGSFGYENTPQFGRAFTAAIV 185
QY 121 DFCIAVKHKLPCLNAAAEELGIPAE-----DSYLAADFLGTCPLSEL 162
DB 186 QCSSSKLEV-NTKICASHGVPPNYPYSPDCLHVDARLFGYDASLAAS-LGKMWAINKP 243
QY 163 QQSRKXVFASMY 173
DB 244 SNSNRATHNLY 254
RESULT 15
AA014800
ID AA014800 standard; Protein; 402 AA.
XX
XX AA014800;
AC
XX
XX 11-DEC-2000 (first entry)
XX
XX Erysipelothrix rhusiopathiae protective polypeptide antigen fragment.
XX
XX Protective polypeptide antigen; PPA; Fujisawa strain; swine erysipelas;
KW recombinant production; immunogenic; subunit vaccine;
XX transmucous administration.
XX
XX Erysipelothrix rhusiopathiae.
OS
XX
XX JP3072345-B1.
XX

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - nucleic search, using frame\_plus\_p2n model  
Run on: November 7, 2003, 20:29:33 : Search time 35:0 Seconds  
(without alignments)  
2307.723 Million cell updates/sec

Title: US-09-613-486-15  
Perfect score: 991  
Sequence: 1 MELMDSNLSNLVITASSL.....GGVATPVSNLRQLGRRVM 198

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n model -DB=xih  
-Q=/cgn2\_1/USPTO\_spool/US09613486/runat\_07112003\_120411\_27035/app\_query.fasta\_1.391  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LCOPEXT=0  
-UNITs=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09613486 @CNC 1.1 3508 @runat\_07112003\_120411\_27095 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=106 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vt.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_cm.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_to.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rnd.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	597	6	AR138306 Sequence
2	991	100.0	15000	14	AF039204 Grapevine
3	991	100.0	15500	6	AR138299 Sequence
4	984	99.3	8590	14	GLAV4131
5	921	92.9	18527	14	AF314061 Grapevine
6	297	30.0	15468	14	AF190581 Beet yell
7	296	29.9	612	14	BYU7-295
8	295	29.8	15468	14	AF056575 Beet yell
9	293	29.6	2724	14	BYVCPG
10	293	29.6	5980	14	BYVMBP
11	286	28.9	6746	14	CLBYV3PH
12	286	28.9	15480	14	BYVCA
13	285	28.8	783	6	A41914
14	266.5	26.9	10545	14	BYU51931
15	130.5	13.2	672	14	AF501867 Citrus tr
16	129.5	13.1	672	14	AF184115 Citrus tr
17	125.5	12.7	669	14	AF342894 Citrus tr
18	125.5	12.7	672	14	AF184114 Citrus tr
19	125.5	12.7	672	14	AF342890 Citrus tr
20	125	12.6	1080	14	CTVCOAT
21	125	12.6	19293	14	AY170468 Citrus tr
22	125	12.6	19296	14	CTU16304
23	124.5	12.6	672	14	AF184117 Citrus tr
24	123.5	12.5	672	14	AF184116 Citrus tr
25	123.5	12.5	672	14	AF220502 Citrus tr
26	123.5	12.5	672	14	AF220504 Citrus tr
27	123	12.4	636	14	CTR297702
28	123	12.4	672	14	AF220503 Citrus tr
29	120.5	12.2	589	14	AF456771 Citrus tr
30	120.5	12.2	669	14	AF342895 Citrus tr
31	120.5	12.2	672	14	AF184118 Citrus tr
32	120.5	12.2	672	14	AF342891 Citrus tr
33	120.5	12.2	672	14	AF342892 Citrus tr
34	120.5	12.2	672	14	AF342893 Citrus tr
35	120.5	12.2	672	14	CTVCOATS
36	120.5	12.2	19226	14	CTU56902
37	120.5	12.2	19259	14	AF260651 Citrus tr
38	120.5	12.2	19259	14	CTV18420
39	118.5	12.0	672	14	AF501869 Citrus tr
40	117.5	11.9	587	14	AF220505 Citrus tr
41	117.5	11.9	672	14	AF339088 Citrus tr
42	116	11.7	672	14	AF184113 Citrus tr
43	115.5	11.7	589	14	AF456773 Citrus tr
44	115.5	11.7	672	14	AF501868 Citrus tr
45	114.5	11.6	672	14	AY190048 Citrus tr

ALIGNMENTS

```
AR138306
LOCUS AR138306 597 bp DNA linear PAT 16-JUN-2003
DEFINITION Sequence 14 from patent US 6197948.
ACCESSION AR138306
VERSION AR138306.1 GI:14479815
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 597)
AUTHORS Zhu,H.-Y., Ling,K.-S. and Gonsalves,D.
TITLE Grapevine leafroll virus (type 2) proteins and their uses
JOURNAL Patent: US 6197948-A 14 36-MAR-2001;
FEATURES
    Location/Qualifiers
        source
            1..597
            /organism="unknown"
        BASE COUNT 165 a 121 c 157 g 154 t
        ORIGIN
            1 95e-97 Length: 597
            2 991.00 Matches: 198
            3 100.00% Conservative: 0
            4 100.00% Mismatches: 0
            5 100.00% Indels: 0
            6 Gaps: 0
            DB:
                Alignment Scores:
                Pred. No.:
                Score:
                Percent Similarity:
                Best Local Similarity:
                Query Match:
                DB:
                    US-09-613-486-15 (1-198) x AR138306 (1-597)
                QY 1 MetGluLeuMetSerAspSerAsnLeuValIleThrAspAlaSerSerLeu 20
                DB 1 ATGGAGTTGATGCCACAGCAACCTTAGCAACCTGGTATTAACCGACGCCCTTAGTCTA 60
                QY 21 AsnGlyValAspLysLeuSerAlaGluValGluLysValLeuValGlnLysGly 40
                DB 61 AATGGTGTGCAGCAAGAGCTTTATCTGCTGAAGTTGAAAAAATGTGTGTCAGAAAGGG 120
                QY 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuValAlaLeuAlaAlaArg 60
                DB 121 GTCCTTAACGAGGATTAAGAAGTGGTGTTCGGTCTACTCTCTACGCATCGCGSCNAGA 180
                QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 90
                DB 181 ACCACGCTCCTTAAGGTTTCAGCGCGCAGATTCAGACGTTATATTTCAATAAGTTTCGGA 240
                QY 81 GluArgAsnValValValThrGlyGlyAspLeuLysLysValLeuAspGlyCysAlaPro 100
                DB 241 GAGAGGAATGTGGTAGTACAGAGGCTGACCTTAAGAGGTTACTCGACGGGTGCGGCT 300
                QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaIleVal 120
                DB 301 CTCACCTAGGTTTCACTAATAAATTAGACGTTTCGGTCTGCTACTTTCACCTGAGGCTTACGTT 360
                QY 121 AspPheCysIleAlaIleAlaIleLysLysLysLeuProGlnLeuAsnAlaAlaAlaGluLeuGly 140
                DB 361 GACTTTTGATCGGTATAGCACAATTAACCCCAACTTACCCCGCGCGCGGGAANTTGGG 420
                QY 141 IleProAlaGluAspSerThrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
                DB 421 ATTCAGCTGAAGATTCGTACTTAGCTGAGATTTTCTGGTACTTGGCCGAGGCTTCT 480
                QY 161 GluLeuGlnSerArgLysMetPheAlaSerMetThrAlaLeuLysThrGluGlyGly 180
                DB 481 GAATTACAGCAAGTAGGAGAGATGTTTCGGCAGTATGTCACCTCTAAACCTCAAGGTGA 540
                QY 181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgArgGluValMet 198
                DB 541 GTGGTAATACACCATGAGCAATCTGCGCTCACCTAGGTAGAGGAAGGAGTATG 594
                RESULT 2
                LOCUS AF019204 15000 bp RNA linear VPL 22-MAY-1998
                DEFINITION Grapevine leafroll-associated virus 2 methyltransferase/helicase
```

```
polyprotein gene, partial cds; and RNA-dependent RNA polymerase, putative transmembrane small hydrophobic protein, 65 kDa chaperone protein, 63 kDa protein, 25 kDa diverged coat protein, 22 kDa coat protein, 19 kDa protein, and 24 kDa protein genes, complete cds.
AF019204
AF019204.1 GI:3123909
Grapevine leafroll-associated virus 2
Grapevine leafroll-associated virus 2
Viruses; ssRNA positive-strand viruses, no DNA stage;
Closteroviridae; closterovirus.
1 (bases 1 to 15000)
Zhu,H.-Y., Ling,K.-S., Gosczyński,D.E., McPerson,J.R. and Gonsalves,D.
Nucleotide sequence and genome organization of grapevine leafroll-associated virus-2 are similar to beet yellows virus, the closterovirus type member
J. Gen. Virol. 79 (Pt 5), 1289-1298 (1998)
98264507
9603345
2 (bases 1 to 15000)
Zhu,H.-Y., Ling,K.-S. and Gonsalves,D.
Direct Submission
Submitted (18-DEC-1997) Plant Pathology, Cornell University, New York State Agricultural Experiment Station, Geneva, NY 14556, USA
Location/Qualifiers
    1..15000
    /organism="Grapevine leafroll-associated virus 2"
    /mol_type="genomic RNA"
    /db_xref="taxon:64003"
    <1..-7423
    /function="replication and unknown functions"
    /note="larger than 277 kDa; ORF1a; contains domains for two papain-like leader proteases, a methyltransferase and a helicase; identified by sequence comparison"
    /codon_start=2
    /product="methyltransferase/helicase polyprotein"
    /protein_id="AAC40855.1"
    /db_xref="GI:3123910"
    /translation="ADVAMRLRYCGGKPLVLMRSVLYPDGRCLAHMRYLCAPYCR PFESDLSGDMPTVARACVEKNFGEACGLAUGYITTSRNVHCDYDSAYRYFR NLSGRGGGSFDTLSVTIKISGLPGKPNIAFGAPLCDIRYFVDPVDSGGIOSS AKTRDAARTVEERAAAGSVQPRQKRDGCGRPSGFSHLVGNVNEVRRVVA AGLLRPVGDMDPFRSFTQAGRLVWRSSRSLVCLYSPSKNLFVADVLPCQSD YAMFSGAAGRPFLVYTRIRYPNGFCYLACHYACAFLLSGFDPK3P1GAFPTAA KLANRVSLGERSLCLNLYGATSRGVCHCYDAKF1KDLKLSAVIACKGQVESVY PSBITPAMQKCTEAVYDRLYGDTLSLKLSTSEKODIDFNQVQSLKKRPIKVPFY MSEATQNSRTFRYPQFELKFSHSHSDHPAAAASLENETJVLRCGNSVSDGGCPJ FHUHQTRRVHVCRLPDGDAQRVVRDLQYSNVLGDDDKILGPRKIDICHPVJ GACDHSSAMVMVYKFGSSCYSHKLSI1KDIIMTPTLYTLTWTGFEFLDGEVVMESLD CETEVHADVMVYKFGSSCYSHKLSI1KDIIMTPTLYTLTWTGFEFLDGEVVMESLD KITKSVSPSISCTKLIRVRANSDVVKIIPFADKRMCLPGYDTIVLDSKFSRV FDYVNCASAVSKTEFWWSFIKSSKSVLIISGKI1HKDVLDLKXVSSFAKMLAS GVSRLASEYLANKSHFSDCSFIATSFVUREKIRNMTLNFNLQVLRKVAAT LVDVLDJSTUESITDPACQVAIEJDELGLRAEAENEKIRNLGASIAAKLASEI LVDIDSKPQKQVNSNADKREVRQGLRGRNGVGFELHFPVDSALRLFYA TDQRIKSVYRFLDSAVSLDYNQNSJLIRVLSEGYSCMFALNGLDLSRVSRA VCARKVATSCANASVAKAMITPAAAVCAMFNCSGFGSGDREKYSIHYRTOVLF DTF1FEDSDVLPTEVLSSALCGAIVTLFSSGSI1SANAFLQITKGSFLVNVNVR VTHGLSTAITDGVIRVFSQIVSHLLVGNITGVAYQSA1AGVVP1LVKKCVS1FIL REDTISGFIKHGISFEFLSSILKFLKGLVDLKSITQGVFDSNKHFKPEATQBAIR T1VMQVPAVADKLSAAGKI1YNNFTSRRTFGDESSSDGACEYFSCDEGEGLK GSSYGSFLAFPSIRIMGARBLIKVHCEFCFKLFEFLSKLHFPRTFVFGKNTDV GVDFV1PTGIVETLSSIEBCDQIEELLGDLKGDASLUTDMNYFEFSDFLASITEP PFAGLRGGSKNITAILIAYANHLPRIVAKCSKRP1LAFAPLSSALTEKEKEVPR KSQULVAIVREYQRF1RSRMALGNNFVKS1FADLLPALMKRVSGSFLASVTRPL RGFYSYMCVSAERREKFFALVCLIGLSLFFVR1VGAKACEELVSAARREYF1KIFLR QKTVLSLNF3CHLPSFDVDSASAGLKGAGMT1FHL1VRLSALLSLGWGLKL LSHNLL1FCFALVDVNVLI1KVLGGLSFFVQ1FSLFAAMLQDPFRVEYSEKLVTA FEPFLKCSPPAPALLAGFECCVANSTVSKTRILLRCFVKMLKLRKGLRADGRLH RQKAVP1SNRVVTDGVERLSVMQCGVEALTELRILELDSAVIEKLNRRNRDNTN DDEFT1PAHEQMOEVTT1FCSKANSAGLAEARLVDAIKSKLSKLVNEMVRKSGST SESEVAVLSDDEAVEE1SVADERDSDPKTRVISEVNLNLSNFSFEFKPI1VVDNDKDTG
```

```

GLNVAEPEYVYMGELALFEIHSKLCITYDCLRIWVDFRVSAPCSDEAQLYVKKQSTRI
VQKEVRLH:KPHEDHDFDQKISINKARGGNVJYHNLAFNLAFNLKFLGAPPSRS
FVTNSVD:LLYEAPGGQKTYDIDSLFLVKFKGSEVMTILTANKSQVEILKXVE
KEYNSICQRKQKSPKSIYITDILMEHRCGDADVLFIDCFMVHAGSVLACIEF
TRCHWIMFDGSRQ:HYIERNELCKJYGLDRFVQLQCRVYGNISYPCPMDVCWMLS
TVGNLIATVKGSGEKGSRNIEKNSVDLVPDVGSTFLCMQSKLBIKSHFTRKG
LTKNLVLTVEAOGELYARVNLVRLAFORDEPEKSIPIHUTVALSRHTDSLT:YVLAAR
RGATCDAIQKAEELNKRFAVFTSPGSGSVINLWKKVDVEDSRCSASSAPLSVNDP
LNEVNPADVDFGDLUSADFTSPFEGGASG:VVRDNISSNITDHCKORV
<7422..8801
/function="replication"
/notes="RdRp; 52 kDa; similar to RNA polymerases of other
c.isteroviruses; presumably expressed via +1 ribosomal
frameshift"
/codon_start=1
/product="RNA-dependent RNA polymerase"
/protein_id="AAC40856.1"
/db_xref="GI:3123912"
/translation="SVVRSQAIPIRRKPSIQENLYSPEARVNFSTCDRTNSMFOEA
YANCLRRQCDLDAFSLRDVLSI:RSIGIECHLEKRTSQ:KALMKQVDSPLEIDCE
-CRFKLVXRDAKVKULDOSSCLTHSAAQNIHFRRKSI:NAIFSPIFNEVKNRMCLKP
NIRKFTMTNRDFASVVSNGDGVHIGEDVFSKYDQSDAFKAFSEVWYKELGV
DELLAIMCGERLSANTLDGQSFTEIENQRKSGASNTWIGNSLVTLGLSLYEDVR
NFALYISGDSILFRSE:SNVADIDICTDMGFETKPMSPVPYCSKFWMCXKHTF
FVDPYKLVK:GAVKEDVSMDFLFTFTFSFKDLSDFENDERLOKLAELVALKYEVQ
TGNLTALSVIHLCSNFLSFKSLYPRKGMQVYTSVKKALKKSCSLFDSFY:TPFG
QAVWWDDE"
8805..9035
/notes="6 kDa; probably membrane-associated; similar to
small hydrophobic proteins of other closteroviruses"
/codon_start=1
/product="putative transmembrane small hydrophobic
protein"
/protein_id="AAC40857.1"
/db_xref="GI:3123912"
/translation="MNQVQFECFLNLAVFAVTFIFILLVFRV:KSPRQKGHEAPV
PVVRGGFSTVV"
9051..10850
/notes="p65; HSP70; similar to heat shock 70 proteins;
identified by sequence comparison"
/codon_start=1
/product="65 kDa chaperone protein"
/protein_id="AAC40858.1"
/db_xref="GI:3123913"
/translation="MVVFGIDFGTFTVCYKQGVKGFSPKQNN:SAVISTYLFPDS
NHTFGYEAESLNLKVKGSFVRDLKRWKVCSSNLDAYDLRLKHYSVRLVK:QSG
LNETVSGNFGVYKNSAHLPG:ALPIKAVICAGFACTCTGVICSVPANVDYQ
RNFCDQCV:SGYQCVYINPEPSAALASCNISGKKSANLAVYDFGGGTFDVS:SYR
NN:FVVAAGSDNLGGRDVRAPLTHLSLTSJEPDLTDISN:KESLSKTDAEIV
TLRGVGRKEDVRNKNILSVMLPYVNRILKIELSTLYAKSMNESARVKDGLVL:
GGSYLPG:ADVLTKHQSVRILRVSDPRAAVAGCALYSCLSGGGLLIDCAAH:
VALADRSCHOI:CAPAGAPIPFGSMPLY:ARVKNKSOREVAVFEGEYVKCPNKKIC
GANTRFIDVGTGDSYAPVTFYMDFSISSVGVSFVVRGEGKQVSLTGTPAYNFSVY
ALGSRVREHL:SLNKVFIQLLJHRKADRRILF:KDEAIRYADS:DIADVLEKYSY
AASAPFDEVEDILLKSVQVLRGSLSE:PL"
10777..12432
/notes="p63; putative heat shock protein 90 homolog"
/codon_start=1
/product="63 kDa protein"
/protein_id="AAC40859.1"
/db_xref="GI:3123914"
/translation="MNSYMSIESIFKKPYGEADWKVLSRIAHSSEIKTLPDRLYG
GRVVKGEFSALESNPEQGLGILFISEREVGSKLCGITVBEAAVDLNPKAYKTA
ETCSPDVKGEOKYSMEDVMNFMPLSLDNDVKVLTPOCKLSNSGCE:INPDKGRF
VALTFKDRDADT:GAANVRGVDYLVYAMS:FEQTKQSCQSGNISLETKCEYRTI
YLG:DLFPTADPILTLGLIDYFCKEYNYFYSSYKRVNFRFFLAMP:SDVF
FVQVKAPDVRLLFQAABALEVLEPTLSIDSVVGHILRYVESY:SDPAIDALE
DKLZAILKSNRPLSLAQWALFFCYGEFTAQSRVORPQVYKTPDQSVGGFEYKVK
DVSEFFKLORELPSLRQCPNGARAHEAFKIPNGN:SFRP:SLKVPREFWYK:
DYFRHANESGI:TEEILILNINSVDVFK:CAERACNLTLPKAKBFSKHNKSNIGSSQE
RRIKDPLVLKDTLYEFQHRKAGSGSRTDLGSRADHAKGSG"
12344..13015
/notes="p25; CPd; coat protein duplicate"

```

```

/codon_start=1
/product="25 kDa diverged coat protein"
/protein_id="AAC40860.1"
/db_xref="GI:3123915"
/translation="MSSNT:SVPGGJEALETSGVLTTRKEAVKFFNELKNENYSV
DSSRLSDSEKVEKESKESFSELSATDBHEVYHIFELIRCAKISTSEKVKYVGGH
TYVDGKTYTVLDRAWFNWMSLTKYKRVNGLPAFCACEDLYLTVAPIMSEFEKTK
AVMGKGPVKGKEYLCAFLSGTSKLMSECHRAVSIVAANKNAVDASAF:GGERKIVSLY
DLGRY"
13080..13680
/notes="p22"
/codon_start=1
/evidence=experimental
/product="22 kDa coat protein"
/protein_id="AAC40861.1"
/db_xref="GI:3123916"
/translation="MELMSPSNLSNLVITPASSINGVDKLLSAEVEKVLVQKGPME
G:EVVFGLLIYALAAARTSPQVQADSV:FSSNFGSRNVVVTGDKKVLIDGCAPI:
RFTNKLRTFGRTFTEAVVDPCIAVKHKLPLONAAAE:GIGPAEDSYLAADFLGTCPKLS
ELQSRKMFASMYALKTEGGVNT:PVSNLRQLGRREV"
13680..14165
/notes="p19"
/codon_start=1
/product="19 kDa protein"
/protein_id="AAC40862.1"
/db_xref="GI:3123917"
/translation="MEDYEKSESLLIRTNLNTMLLVKSDASVELPKLLICSLRV
SGRGEVCCNREELTRDFEGNHHTVIRSRIOYDSESAFEFNNSDCVVKPFJSTGSV
FWFRLSETKGRAVRH:RTFPEANRFFPGHCGTMEYCKQVLTETES:IDSFCEERN
R"
14167..14784
/notes="p24"
/codon_start=1
/product="24 kDa protein"
/protein_id="AAC40863.1"
/db_xref="GI:3123918"
/translation="MRVIVSPYEAEILKRSTDMRLNIDSGVLSKCEIKAFSTITRD
LHCASVOMGVDTGLYORNCAEKRLIDTVESNIRLAQPLVREKVAHFCKDEPKELV
AFIRTKVYGLGVREAVKEMBSLAKTKVLNKMSEMAFYMSPRAMKNAEWLELKFS
PVKFRDLLLVGYELNELCAEDDVHVVKVNGENGDHLELDQDSC"
14785..15000
/notes="similar to 3'UTRs of other monopartite
closteroviruses"
BASE COUNT 3877 a 3050 c 3820 g 4253 t
OR:GIN
Alignment Scores:
Pred. No.: 1.37e-95 Length: 15000
Score: 991.00 Matches: 198
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Idels: 0
DB: 14 Gaps: 0
US-C9-613-486-15 (1-198) x APO39204 (1-15000)
QY 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20
Db 13084 ATGGAGTGTGATGTCGACGACGACCTTAGCAACCTGGTATACCGACGCTCTAGTCTTA 13143
QY 21 AenglyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40
Db 13144 AATGGTGTGACAGAGAGCTTTATCTCTCAAGTTGAAAAAATGTTGGTGCAGAAAGG 13203
QY 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyAlaLeuAlaAlaArg 60
Db 13204 GCTCCTAAACGAGGCTATAGAGAGTGCTGTCTCGTCTACTCTTACGACTCGCGCAAGA 13263
QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
Db 13264 ACCACGTCTCTCAAGGTTACGCGCGACGATTCAGACGCTATATTTCAAATAGTTTCGGA 13323
QY 81 GluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaPro 100

```

```

13324 GAGAGAAATGTTGATGACAGAGGGTGACCTTAAGAAGGTACTCAGCGGGGTGCGGCT 13383
101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyVal 120
13384 CTCACCTAGGTTCACTAATAAAGCTTAGAACGTTGGTGGTACTTTCACTGAGGCTTACGTT 13943
121 AspPheCysIleAlaTyValLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140
13944 GACTTTTGATGTCGATATAGACAAATTAACCCCACTCAACCGCGCGGGAATGGGG 13503
141 IleProAlaGluAspSerTyValLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
13504 ATTCAGAGTGAAGATTCGTACTTGTGTCAGATTTTCTGGGTACTTTGCCCGAAGCTCTCT 13563
161 GluLeuGlnSerArgLysMetPheAlaSerMetTyValAlaLeuLysThrGluGlyGly 180
13564 GAATTACAGCAAAAGTAGGAAGATGTTCCGCGAGTATGTACGCTCTAAAACTGAAGGTGA 13623
181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgThrGluValMet 198
13624 GTGGTAAATACACCAAGTACCAATCTGCGTCAGCTAGGTAGAGGGAAGTTATG 13677

RESULT 3
LOCUS AR138299
DEFINITION Sequence 1 from patent US 6,197,948.
ACCESSION AR138299
VERSION AR138299.1
KEYWORDS GI:14479808
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1. (bases 1 to 15500)
AUTHORS Zhu, H.-Y., Ling, K.-S. and Gonsalves, D.
TITLE Grapevine leafroll virus (type 2) proteins and their uses
JOURNAL Patent: US 6,197,948-A 1 36-MAR-2001
FEATURES
    source 1..15500
    /organism="unknown"

BASE COUNT 3984 a 3178 c 3970 g 4368 t
ORIGIN

Alignment Scores:
Pred. No.: 1,44e-95 Length: 15500
Score: 991.00 Matches: 193
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-613-486-15 (1-198) x AR138299 (1-15500)

Cy 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20
Db 13584 ATGGAGTTGATGTCGACACCAACCTTAGCAACCTGCTGATACCGACGCTCTAGTCTA 13643
21 AsnGlyValAspLysLysLeuLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40
13644 AATGGTGTCCACCAAGAAGCTTTTATCTGCTGAAGTTGAAAAGATGTTGTCGAGAAGGG 13703
41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyValAlaLeuAlaArg 60
13704 GCTCTAACGAGGGTATAGAAGTGGTCTCGGTCCTACTCTTACGCACTCGCGGCAAGA 13763
61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
13764 ACCAGCTCTCTTAAGGTTACGCGCGAGATTCAGACGTTATATATTTCAATAGTTTCGGA 13923
81 GluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaPro 100
13824 GAGAGGAAATGTGGTAGTACAGAGGGTGACCTTAAGAAGGTACTTCCAGCGGGTGTGCGCT 13883

```

```

101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyVal 120
13884 CTCACCTAGGTTCACTAATAAAGCTTAGAACGTTGGTGGTACTTTCACTGAGGCTTACGTT 13943
121 AspPheCysIleAlaTyValLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140
13944 GACTTTTGATGTCGATATAGACAAATTAACCCCACTCAACCGCGCGGGAATGGGG 14003
141 IleProAlaGluAspSerTyValLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
14004 ATTCAGAGTGAAGATTCGTACTTGTGTCAGATTTTCTGGGTACTTTGCCCGAAGCTCTCT 14063
161 GluLeuGlnSerArgLysMetPheAlaSerMetTyValAlaLeuLysThrGluGlyGly 180
14064 GAATTACAGCAAAAGTAGGAAGATGTTCCGCGAGTATGTACGCTCTAAAACTGAAGGTGA 14123
181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgThrGluValMet 198
14124 GTGGTAAATACACCAAGTACCAATCTGCGTCAGCTAGGTAGAGGGAAGTTATG 14177

RESULT 4
LOCUS GLAV4131
DEFINITION Grapevine leafroll-associated virus 2 genes encoding RNA polymerase and coat protein, hsp70, hsp90 gene and ORF2, ORF7 and CRF5.
ACCESSION Y14131
VERSION Y14131.1
KEYWORDS GI:2765391
SOURCE coat protein; heat shock protein 70-like; ORF1; ORF2; CRF3; ORF4; ORF5; CRF6; ORF7; ORF8; RNA polymerase.
ORGANISM Grapevine leafroll-associated virus 2
REFERENCE 1. (bases 1 to 8590)
AUTHORS Abou-Ghanem, N.
TITLE The nucleotide sequence of the 3' terminal region of grapevine leafroll associated closterovirus 2
JOURNAL Unpublished
AUTHORS Abou-Ghanem, N.
TITLE Direct Submission
AUTHORS Abou-Ghanem, N.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-1998): N. Abou-Ghanem, Universitat degli Studi di Bari, Dipartimento di Protezione delle Piante and Centro Studio virus e virosi delle colture mediterranee, via Amendola 165/A, I 70126 Bari, ITALY
REMARK 3 (bases 1 to 8590)
REFERENCE 3 (bases 1 to 8590)
AUTHORS Abou-Ghanem, N.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-1998): N. Abou-Ghanem, Universitat degli Studi di Bari, Dipartimento di Protezione delle Piante and Centro Studio virus e virosi delle colture mediterranee, via Amendola 165/A, I 70126 Bari, ITALY
COMMENT On Jan 10, 1998 this sequence version replaced gi:2369864.
FEATURES
    Location/Qualifiers
        1..8590
            /organism="Grapevine leafroll-associated virus 2"
            /mol_type="genomic RNA"
            /db_xref="taxon:64003"
            /lab_host="N.benthamiana"
            1165..2427
            /note="ORF1"
            /codon_start=1
            /product="RNA polymerase"
            /protein_id="CAA74561.1"
            /db_xref="GI:2369865"
            /db_xref="SP:RMBL:O39852"
            /translaton="MFGEAMMNCRLRCFOLDAFSSLRNDVISTRSIGIEQWLKGRTP
            SR:KALMKDVESPLEIDDEICRFKLMVKRDARVKLDSCLTKHSAQINMFHRSKINA
            IFSPIFNEVKRIMCCLKPNIKFTFTMTNRDFAVSVMMLGDDDDVYHIGVDFSKYDK
            SQDAVKAFEVYKELGVDBELLAIWMCGERLSIANTLDGOLFSTIENQKSGASNT
            WIGNSLVTLGLSLYIDRNFEALYISGDSLIYSRSEISNYSADDICTDNGFETKFN
            PSVPYFCSKFVWCGHKTFFVLDPYKLVKLGAVKEDVSMDSLFTFTSPKDLTSDFN

```

```
CDS
DERLIQKAEVLAKYEVQNTTTLASVHICLRNLFQSFSLYPRVKGQVFFYSVK
KALLKSCGSLPDSFV:PFQAVVMWDE"
2491..2661
/notes="ORF2"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAA74562.1"
/db_xref="GI:2369866"
/db_xref="SPTREMBL:O39853"
/translation="MNQVLOFELFLNLAVFAVTF:FI::VFRVIKSPRQKGHEAPV
PVVRGGGFSVW"
2677..4476
/genes="hsp70"
2677..4476
/genes="hsp70"
/codon_start=1
/product="heat shock protein 70-like"
/protein_id="CAA74563.1"
/db_xref="GI:2369867"
/db_xref="SPTREMBL:O39854"
/translation="MVVFLDGFCTFTSTCVYKGRVSEFKQNSAYIPTLYL:ESDS
NMHTFGYEASJSLNKLKSGSFYRDLKRVGCDSSNLDAYLDR:LKPHYSVRVLK:GSG
LNTVTSIGNPGTVKSEAHLPGLIALFKAVISCREGAFACCTGTCVICSVPANYDSVQ
RNTFDQCSUSG:CCVY:NEPAAALSACNSVGKSKAN:JAVYDFGGG:FDVSIISYR
NTFTVRASGGD:NLGG:RDVDRFLTHLFSLTSLDPDLTVYSLNKES:SKTDAEIVY
TERGVDRKEDVRVKNILTSVMLPIVNR:TLKILESTLKYAKSMNESALVXCDLV:J
GSSYLPGLADVLTHQSVDRILURSDPRAAVGALYSSCLSGSGG:LLIDCAHT
VA:ADSSCHQIICAPAGAPIPSSGMPLYLARVKNKSOREIAVFEVEYVKCPKNKIC
GAN:RFEDIGVTCDSVPTFFYMDFSISSVGAVSFVVRGPEGQVSLTCT:PAYNFSSY
ALGSRFVREHJSLNNKVFUGLJLHRAKDRRLITKDEAIRVADSI:DIADVLKEYSY
AASALPDEDVEILLKGSQVKVLGRSRLEEIPL"
4403..6058
/genes="hsp90"
4403..6058
/genes="hsp90"
/codon_start=1
/product="heat shock protein 90-like"
/protein_id="CAA74564.1"
/db_xref="GI:2369868"
/db_xref="SPTREMBL:O39855"
/translation="MNSYSHESL:PKPYGEADHKKLYLSRIAHSSEIK:TLPDIRLYG
GRVVKSEFSA:LPNFEQLGUFILSEREVGSKLGGITVEAAAYD:JNPKAYKPTA
ETSPVPDGGQGYKNEVDNMNRLNLDVNDKYLAEQCKSLNSGSEL:YNDDKGRF
VALTFKDRDPTAGTANEDVNLVYVAMSLEFQRTQKSGSNGISLYEKYCEVIR
YLSGTLFFTPAPDRIPLETHIDFCKEYNIFYSYKRVNDNFRFLAKYMLP:LSVF
FQWQVAPADVPRLFLLSAEALT:EVP:TLSDISQVVGHLISYVESY:SDPAIDALE
DKLEAILKSNPRLS:TAQ:LVGFCYGYGEPRTAQSVVRBPYKTPDSVGFEINMK
DVEKFKQREDPVS:LRQCFNGARAHAFAKIFKNISFRPISRLNVPREFYLN:
DYPFRANRS:GTBEEIILINISVDVRKJCAESACSTLPSAKRFSKNHKSNISQSRQE
RRIKPLVLVKDTLYEFQHKRAGMSGRSTRDLGSRADHAKGS"
5970..6641
/notes="ORF5"
/duplicate_gene"
/codon_start=1
/product="coat protein"
/protein_id="CAA74565.1"
/db_xref="GI:2369869"
/db_xref="SPTREMBL:O39856"
/translation="MGSNTSPVPGLEALETSGVLLTTRKEAVDKFEFNLKNEYSSV
DSRLSDSEKVEKLEKSKSPKSELASTDEHFYHIIFFLRCAKILSTSEKRYVGSX
TYVVDKTYITLVDANFNWMSKSTKKIKRVNGURAFCCACEDLYLT:VAP:MSERFKT
AVGMKGLPVQKEYLGA:DF:SGTSKLSMDHRAVSIVAANKNADRS:FTGGERKIVSLY
DLGRY"
6710..7306
/notes="ORF6"
/codon_start=1
/product="coat protein"
/protein_id="CAA74566.1"
/db_xref="GI:2369870"
/db_xref="SPTREMBL:O39857"
/translation="MLMSDSNLSNLVITDASS:NGVDKKLLSAEYVVKLVKQKAPNE
GIEVFGLLLYALAAARTSPKVRADSDVIFSNSPGRNVVVTGGLKKVLGGCAPLT
RFTNKLRITFGRITFEAYVDFCIAKYHKL:POLNNAAEGLIP:PAEDSVLAADFLGT:CPKLS
```

```
CDS
ELQSRKXKPSVYALKTEGGVVNTFVSNLRQLGRREVW"
7306..7793
/notes="ORF7"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAA74567.1"
/db_xref="GI:2369871"
/db_xref="SPTREMBL:O39858"
/translation="MEDYEKSESILILRTNLNTML:JVVKSDASVELPKLILCGLYLV
SGFEVTCNREELTRFEGNHHIVIRIIQYDSESAFAEFNKSDDCVVAKFLFTGSV
WFFELRSTKGRAVRLHRLTFEANNFFFGSHCGTMEYCLKQVLTETESIIDSFCEER
R"
7793..8410
/notes="ORF8"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAA74568.1"
/db_xref="GI:2369872"
/db_xref="SPTREMBL:O39859"
/translation="MRVIVSPVEADILKRTDMLRNI:DSGVLSTKECI:KAPSTITRD
LHCASQWQGVDTGLVQRNCAEXRLIDTVESNIRLAQPLVREKNAVHFCKBPKELY
AFITRKVELTGVGNRAVREKMRSLTKLVNKNLSLEAFYMSPRAMNAEWLELKFS
PVKIFRDLJLDVETNELCAEDDVHVQKVNENEEDNHDLELCQRC"
BASE COUNT 2344 a 1686 c 2135 g 2425 t
ORIGIN
Alignment Scores:
Pred. No.: 3,73e-95 Length: 8590
Score: 984.00 Matches: 197
Percent Similarity: 99.49% Conservative: C
Best Local Similarity: 99.49% Mismatches: 1
Query Match: 99.29% Indels: C
DB: 14 Gaps: 0
US-09-613-486-15 (1-198) x GLAV4131 (1-8590)
QY 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20
Db 67.0 ATGGAGTTGATCTCCGACAGCAACTTAGCAACCTGGTGATACCGACGCTCTAGTCTA 6769
QY 21 AsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40
Db 6770 AATGGTGTCGACAGAGAGCTTTATCTCTCTGAAGTGTGAAAAATGTTGGTCAGAAAGG 6829
QY 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArg 60
Db 6830 GCTCCTAACGAGGGTATAGAGTGGTGT:CGGCTACCTCCCTTACGCACTCGCGGCAAG 6889
QY 61 ThrThrSerProLysVal:Gln:ArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
Db 6890 ACCACGCTCTCTAGGTT:CAAGCGGCAGATTACAGCGTTATATTTCAATAGTTTCGGA 6949
QY 81 G:uArgAsnValValVal:ThrGluGlyAspLeuLysValLeuAspGlyCysAlaPro 100
Db 6950 GAGAGCAATG:GGTAGTAACAGAGGGTGACCTTAAAGAGGTACTCGACGGGTGTCGGCT 7009
QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrVal 120
Db 7010 CTCACATAGTTCACTAATAACTTAGAACGTTCCGTCGTACTTTTCTACCTGAGGCTTACGTT 7069
QY 121 AspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaAlaGluLeuGly 140
Db 7070 GACTTTTGTATCGGTATAAGCACAAATTAACCCCAACTCAACGCGCGCGGAATTGGGG 7129
QY 141 IleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
Db 7130 ATTCCAGCTGAAGATTCGTACTTAGTCAGATTTTCTGGGTACTTGGCCGAAGCTCTCT 7189
QY 161 GluLeuGlnGlnSerArg:LysMet:PheAlaSerMetTyrAlaLeuLysThrGluGly 180
Db 7190 GAATTACAGCAAGTAGGAAAATGTTCCGAGTATGATCGCTCTATAAAACCTGAAGGTGGA 7249
QY 181 ValValAsrThrProValSerAsnLeuArgGlnLeuGly:ArgArgGluValMet 198
```

250 GTGGTAATACACCTGACGATCTCGCTACCTAGCTAGGAGGAGGAGTTAG 7303

AF314061 16527 bp ss-RNA linear VRL 06-NOV-2002

DEFINITION Grapevine rootstock stem lesion associated virus methyltransferase/helicase polyprotein, RNA-dependent RNA polymerase, p6, Hsp70-like protein, Hsp90-like protein, coat protein duplicate, coat protein, p19, and p24 genes, complete cds.

AF314061

AF314061.1 GI:24636914

VERSION 1 (bases 1 to 16527)

KEYWORDS Grapevine rootstock stem lesion associated virus

SOURCE Grapevine rootstock stem lesion associated virus

ORGANISM Viruses; sakna positive-strand viruses, no DNA stage; Closteroviridae; Closterovirus.

REFERENCE 1 (bases 1 to 16527)

AUTHORS Zhang, Y. and Rowhani, A.

TITLE Nucleotide sequence of grapevine rootstock stem lesion associated virus

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 16527)

AUTHORS Zhang, Y. and Rowhani, A.

TITLE Direct Submission

JOURNAL Submitted (16-OCT-2000) Plant Pathology, University of California, One Shields Ave., Davis, CA 95616, USA

FEATURES

Location/Qualifiers

1..16527

/organism="Grapevine rootstock stem lesion associated virus"

/mol\_type="genomic RNA"

/db\_xref="taxon:167634"

107..8902

/note="327 kDa"

/codon\_start=1

/product="methyltransferase/helicase polyprotein"

/protein\_id="AA634615"

/db\_xref="GI:24636915"

/translation="MSSLAISALPSAVAOLSFQOPATVAKSFYMTSLPS:CTCFQOC

EFSSFLIC:ASFQK:KTFISF:SSVRELFALFSELSTGFCCEFI:GLGGEHFLFPVP

QCGLPSVCLGVNG:FAVLN:SPFF:ADVASJLMDVGVVTPFKVR:SKRAKLNKRAF

STARAAAR:RRKVGNGKEVRI:SHVQF:PAAGLFQSNKNGTELII:IGRDVGVARVKT

AAPVRRGGGGRVPLIP:VAVSPFGVRYCAPSCGAIACPKNLASRAANDSRAE

KILSARRKLQOONFNSVAKAROVTAASRVMAQGEVTPPHRI:FGFSFEPFP:PS

SGMARSV:REEPALV:CGASAPAIK:ETPVS:VKKAD:SPAAPV:PEALTPRAVKEGA

IVLSGVHEVLANK:REGVLR:PRVGGDMRPSR:FSSTHGC:KIYANVR:NTSWLACTKV

NEKFSVA:PCSDVY:VAMLY:VSGGK:FPVL:KSRCEY:PDGLCI:MAHMY:FLCAFFCQFP

RESDYPLGSPSVARLUKALVKRFGDEALNIGVGYVSRRAFPDGYDYSKYARS:LMRL

AGYVGDEVT:HLSTL:VTLKPGERADR:PKS:SEGSFOCE:RYVESVNPAPR:LRSE

KVLRKVPENGVHSTESADVGSNVRKRDARIDRFKSHLLAGSNVKVKR:LAAGV:LYR

VGGMDP:PHNSFL:TOAGV:HLVRRKTS:SSVCVELSTPDGLLRDVI:PCSRDYAALSF

AQGRPEP:VLMTRDKYKNQCYLAHCYAS:ALLKGFH:PAVFCIGANPTAK:LSRMV

SULGDRS:LINLYG:STSGI:FHCDDAAYVKDLRFMSAIVAKEMGEVNASD:TPA

XKQTEATV:R:JGNSDLSL:KUSIEKO:IDFNKDVQSLKRD:KIRPVFFVNSVETQN

SUTRFYQELFELKSTESHDHFAAASRLLENETVRLGCHSVSD:GSCCLFHLNKSX

QHWVHRVJDLQKQRRVRLQYSSVRSQDEKILEGPNVNDVHPYLCQCHES

STMMVQVYDAS:YEICSA:MIKKKARITV:TWVTFGE:JDKGECVMS:LDQIEVDV

HADVMYKFGSSCYSHKLSI:IKDITTPV:VLGGGLF:SVEMVEIRMGVNYFKITKSEV

CNIVSKLLRYRANSDVUKVLP:FPDKKRMCLPGV:II:VLDSKFVS:RVPDYVVCN

CSAVNSKTEWNSFKSKSRV:ISGKIIHKDNLDLKIVSFESFAVWLASGRSLA

SEYLANLGHFGSCGCFIATSFVLNBEIRNMTLNFNECJLGLVKRVAFATLDVDFD

LDATMESITDPAECVAVDLBELGLCRLANAENEKTRNLAGDSMAAKLASEI:VVDIEAF

VPVAAHEHPKSEVEKKEHAGLGRGSKAGVLKEFLDFVLSASRLPSCVADQQR:IK

SVYHLYLQTVSPVDQNVENLSPFTMRVLSKGYSCMPSYLANRGLSARVASA:VQIVREI

AVLCANVUSKAKAVVYTSAAVCAAMFNSGSGGGRYKSYMHRYTQVLEDTI:FPED

SATYLP:ELLSA:ICGAITV:TSAGTSVSNALFLLQITKFSLELIVRNVRVTHSLST

TATDGMVRSVFSQI:SHLLVGNNTLAYQIAFLTVPLLVKCVGLI:FMSSDEITYSG

FIKHGVSESPITNLYKRGVLDELKAVQGVDFDSNKHLPDRTQTAIRST:WQIIP

TAVVDGLKCAASVSGVIGVKNTNVEKSGGGDDEHYFSCDESEYFGLRGSGRS

FSILFFESFVLRPAR:IT:IGRI:KHRCFSK:IESI:SLK:HE:RTFLLGKXPSDVGYEFL

SEDDVALSISIECQSIENLIGGFRDMADDSLTQNYDFDSDEFLASVNGLP:CAGLR

GGGRVGMALAILNVYNSVINSIMARAASKNSTFTTSLIFSRGLVDLRINIFS:KESVSE

LVRYKYVNF:KSRMRGLRLNHEHVNSFGELVPLLLRNKLSETFI:IRAVPMSKWCKV

CLSAERERFFALVCLVLTGFPYFARFLGGVCDKLLSLSRAYDRVIVFLKERCDA

MNGLRCLFHEVDNGDSTAGLKGGSQTLWGLLRLVRFALPSLGVCKCFVSHRDL

FF:CVSLFHEVENVMIKILGCLSFSSRPISALAFALLOPERFVEYSGKSLTFLPFLK

ISPRVPLKGGFPPEVAGSP:LLRGVRRFLRI:IAVKLNLRKRLGRTRNGGYSQRAVP

VIFSNVTKDGCERLSEKMGVEALRSRLRI:FEELDAAVVEKLNKRNREADDSERSR

PMHESHTDVTSPGSKANSVGTALERAVLVEDAVKSEKISMAKDLVRKCTVSDDE:TN

TLSDDEAIEBVPVSEERDSPKTVRVSEVLELNRLNSFEFPKPIVDNDKNETGLINAV

REFFYMOELALFEIHSKLCAYDQLRI:VNFERSMAPCNEAOQIVYKNGSTIVQGEV

RSNIKFDHDFLFDGKISVNRNRRSGNVLVYHNDLSFLASNLFLAGYFSRSFVFN

SVDLIYEAEPGGGKTTIIDLSPFKVKGGEVSTMLTANKSKSQVEILAKKVEDISNI

ECOKRDKTPKKSIVTIDAYLWHRGCDILFIDECFMVHAGSVACIEFTRCHKV

MEFDSRQIHYIERNELDKLYGDL:DRFYDLOCRVYGVNSYPCMPWYCAWLSVYDVM

IATVGESEKSSMRINEINSVDDLVFDMGSTYLCMLQSEKLEISKHFTRKGLSKNV

LTVHEAQGETYARVNLVRFLQDEDFPKSIRH:TVALSRRHTSLTITVNLAAARGATC

DA:OKANELVNFERVFT:FGGGSVINDVKVDVEDNSRCKKASSAPLVS:NDF:JSDVNP

GT:VIDFGDLSADFTGPFECGASGIVVRDNISSNITDHDQKV"

/note="53 kDa; translated via ribosomal frameshift"

/codon\_start=1

/product="RNA-dependent RNA polymerase"

/protein\_id="AA63467.1"

/db\_xref="GI:24636916"

/translation="SAVRSCAIPRRKPSLQENLYSPEARNYNFTCDRTFSASVFGEA

MANNCIRRFDELTFFSLRDDVISITRSGTEQWLEKRTPSQIKALMKDIESPE:ODE

ICRPKTMKXDAKVDLSCLTKHSPQAMIMFHRKSINAI:FSPIFNEVKNR:MSCLKRP

NIKFPTMNRDFAVUSNMGLDGDYVHIGEVDFSKYDKSQDAFAVKSFEVNYKELGV

DEELLAIWMCGERLSIANTLDGQLSTYTIENQRKSGASN:TWIGNSLVTLGILSL:YDVR

HFAPLVSGDSDSLIFSRNKISNYADDICADMGFETKFPSPSVFYFCSFVWMCCHKTF

FVPDPYKFLVKAGKEDVMDLFETF:SFKDLTSDFNDETV:QKLAEVALKCYDVQ

SGNTALALSIV:HCLRSNFI:FSKLYPRVKGMQVY:SVKKALFKSGCSLYSVTPFG

HAVWVWDE"

10373..10543

/note="6 kDa"

/codon\_start=1

/product="p6"

/protein\_id="AA63468.1"

/db\_xref="GI:24636917"

/translation="MQVQLQFECLYLNLQAVPAITFVFILLIIRV:KSFERNVHETPI

SAVGGGQGFSTV"

10552..12351

/note="65 kDa"

/codon\_start=1

/product="HSP70-like protein"

/protein\_id="AA63469.1"

/db\_xref="GI:24636918"

/translation="VVVFGLDGCTTSTVCYVKDGVYKQNSAIVITLYLPSFT

NHWVFGTEASLHMNAKRGSPFRDLKRVGGSSNFSDFDLKPKHYAVRPYKIG:G

LNDTVIIGNYGGTVRSEVHLPLGLIASFIKAMVSCAENAFSCTCTGVCEANVDSVQ

RNFTDQCVSLSGYQCYVMNPESSAALSACNLINKKSANLAVYDFGGCTFDVSIISYR

NNTFVAVASGGLNLGGRDVDRFTLTH:LSLTSLPELSD:SNLKESLSKTA:EV

TLKSGDDKEDVAVNKSIL:TTVMVLYVNRNTNI:EAATLSYAKNMGSLKVKCOLVLI

GGSSYLPLGLDILSKHRRIDRLKVADPAAVAVGALYSSCLSGSGLLILVDCASHT

IATAGSCDQILICAPAGAP:PFSGT:PLYLPAKRNRSQRC:AVFESEYKCPNRKIC

GSNIKEDIGVTSVSPITFYLDV:ISSVGAISFVRGPEGKEVSLTGTPAFTSSV

VLGTRSVREHIS:SNKHVGLL:HKRADRL:LFVSSEGAIVRYVETVADV:KEPFSY

NASLSLPDEVEPELNGKSVQKVLGRGRLEBIPL"

12278..13882

/note="62 kDa"

/codon\_start=1

/product="HSP90-like protein"

/protein\_id="AA63470.1"

/db\_xref="GI:24636919"

/translation="MSNHSWGLSKFYKYGEADWRKYLKSI:VAHTANIKTLPDVQLYG

GRVTKSEFESATPSRPEGLGLMLSERDQWCKMCGITVEAMVQITNP:EAAPLTA

EKCDPVKEGQYKTYMEDVNNFNRNLDINDKILAEQWSSLSNSGVLINPDKBERF

VALTFKNNELVNDSSNANIDCRVDYLVAMS:LFNCRSQKSOAGNISLKEKYCAYIRI

YLENTQILYFSDP:RIPLLTGLYDFCKEYNIFYSTYKRNVDNFRFETLNYMPLISDI

FQWPKPATDHRLLPDISAAELTLEVTYLSVDQSVVIGHVLFVRESYTAIDALE

EKLDAKNSPNHLSIAQWLVGFCYGYGERTAQORVORPGVYKTPNSVGGFEINMK

NVPEFDKIQOREVNPVSLRQFNQGAHAFAKVPKKNISFPKISRLNIPREFWYLV

DYFPHANRSGBSEEBVLNNISVDRVKUACAERACSTLFSAKRFSKSHKSSVPSLQKE

KNYEDPLIALRLNIFRHNRRGRV"



```

13945..14516
/note="25 kDa"
/codon_start=1
/product="coat protein duplicate"
/protein_id="AAN63472.1"
/db_xref="GI:24636920"
/translation="MSSSGTGIAGVGSSEAIETGFLVSTRKEAVDKFFNELRKNENYSV
DSSRISDQKVEKLEKSEKSFSELAETDEHFVYHIFFLIRSAIISTSDKVYGVSSH
KVVVEGKLTIVQGVVWVNNIKGISTKQYKRVNGLRAPCCACEDLYLTVADESERPKTK
ATMGKLPVGVGLGADFLSGTSKLMSDHDRAVSI VAAKNAVDRAFTGGERKIYSLY
DLGRN"

CDS
14604..15200
/note="22 kDa"
/codon_start=1
/product="coat protein"
/protein_id="AAN63472.1"
/db_xref="GI:24636921"
/translation="WELMSDNLGLVITDASSNGVDKKLSAEVIRKLVKQKGAQPSQ
GIETVGLGLYALAAATSTPKVQRADSDIIFONTYQKTVVVTGDLKKVLEGCAPLN
RTNKLRLTGRTFTYAVDFPCVAYKHKFQLNAAAEGLGIPAEDSYLAADFLGACPKLS
ELQQRKMFASMYAKTEGGVNTPVSNLRQLGRREV"

CDS
15200..15685
/note="19 kDa"
/codon_start=1
/product="pi9"
/protein_id="AAN63473.1"
/db_xref="GI:24636922"
/translation="MEICRAQPEALVLRNTONTLLVVSDEVNLPKLLICGVLRV
SHRGDVTGGRNREVDIKDFGAHIVIRSTVYDESAYKEVNNACDGVKFFETGDDV
PFFFQSDIIGRAARHLRIFFEANFFFGSHGCTMEYCLKQVLIETESVIESFCEERN
R"

CDS
15687..16304
/note="24 kDa"
/codon_start=1
/product="p24"
/protein_id="AAN63474.1"
/db_xref="GI:24636923"
/translation="MEKIVSPVEAEEDILTRSTEMLRNIDSGALDTRECIFKAFSTLTDR
LHCASQYQWVDITQCDQCAEKRLIDTVESNLRHAHLIRKVAHFCKDPEKELY
AFITRKYVKLGTVAAEAVKISMRTLSKVLAKLSLEMAFYMSPRAWKAEWLRLRYT
PKIIRDLLLDVDTLNALCEBDEIHVVVTKDEENHDLQEDC"

BASE COUNT      4250 a  3429 c  4224 g  4624 t
ORIGIN

Alignment Scores:
Pred. No.:      5,26e-88      Length:      16527
Score:          921.00        Matches:      180
Percent Similarity: 95.96%    Conservative: 10
Best Local Similarity: 90.91% Mismatches:      8
Query Match:      92.94%     Indels:       0
DB:              14          Gaps:       0

US-09-613-486-15 (1-198) x AF314061 (1-16527)

QY      1 MetGluLeuMetSerAspSerAsnLeuValIleThrAspAlaSerSerLeu 20
Db      14604 ATGGAGTTGATGTCGCATGATACTTGAGCGGTCCTCGTCATACCGACGCTTCTAGTTG -4663

QY      21 AsnGlyValAspLysLysLeuSerAlaGluValGluLysValLeuValGlnLysGly 40
Db      14664 AATGSGTGTGCATTAAGAACTGCTGTCTGCGGAGGTGCATAAAAATGCTTGTGTCAGAAAGGG -4722

QY      41 A.aProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaAsn 60
Db      14724 GGCCCCAGCCAAAGTATTGAACAGCTTTITGGGCTACTGCTCTACCGCTCTCGCAGGCGAGA -4788

QY      61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
Db      14784 ACCACATCGCCCAAGTTCAGAGCGGTGATTCTGATATAATCTTTCAAACACTTACGGT -4844

QY      81 GluArgAsnValValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaPro 100
Db      14844 GACAAGACTGTTTCGTGTGACAGAAAGTGATCTCAAGAAAGTGTTGAAGGTCGCGACCT -4903

```

CY	101	LecThrArgPheThrAsnLysLeuArgThrPheGlyA:GThrPheThrGluAlaTy:Val	120
Db	14904	CTTAATAGATTCACTAACAAGCTGGCAACTTTGGCGCCGACGTTTACGAGGCGGTATGTT	14963
CY	121	AspPheCysLleAlaTyLysHisLys:LeuProGlnLeuAsnAlaAlaAglLleuGly	140
Db	4964	GATTTCCTGGCTGGCTTACAAGCATAGATGCTCAGCTCAACGCGCGCCAGAAATTTGGG	15023
CY	141	LlePrcAlaGluAAsPseTyrLeuAlaAlaAspPheLeuG:YThrCys:roLysLeuSer	160
Db	15024	ATACAGCAGAGAATTCGTACCTAGCTCGGACATTTTATAGTGCTTGCCCGAACTCTCT	15083
CY	161	GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGly	180
Db	15084	GAATTACAAACAAGTAGAAAAATGTTTCCAGTATGTACGCTTTTGAACACTGAGGCGGA	15143
CY	181	ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgGluValMet	198
Db	15144	GTAGTAATACCCCGCTTAGTAATTTCGCCCAATTAGGTAGAGAGAAGTTATG	15197
RESULT 6			
AF190581			
LOCUS	AF190581	15468 bp	RNA linear VRL 04-JAN-2000
DEFINITION	Beet yellows virus strain BYV-4, complete genome.		
ACCESSION	AF190581		
VERSION	AF190581.1	GI:6492367	
KEYWORDS			
SOURCE	Beet yellows virus		
ORGANISM	Beet yellows virus		
REFERENCE	Viruses: ssRNA positive-strand viruses, no DNA stage;		
AUTHORS	1. (bases 1 to 15468)		
TITLE	Peremyslov,V.V., Hagdiwara,Y. and Dojla,V.V.		
JOURNAL	HSP70 homolog functions in cell-to-cell movement of a plant virus		
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 96 (26), 14771-14776 (1999)		
PUBMED	20079557		
REFERENCE	10611298		
AUTHORS	2. (bases 1 to 15468)		
TITLE	Peremyslov,V.V., Hagdiwara,Y., Alzhanova,D. and Dolja,V.V.		
JOURNAL	Direct Submission		
FEATURES	Submitted (27-SEP-1999) Botany and Plant Pathology, Oregon State University, 2082 Cordley Hall, Corvallis, OR 97331, USA		
source	Location/Qualifiers		
	1. .15468		
	/organism="Beet yellows virus"		
	/mol_type="genomic RNA"		
	/strain="BYV-4"		
	/db_xref="taxon:12161"		
	/note="Californian strain; RNA transcript derived from this cDNA clone causes formation of red necrotic lesions in Claytonia perfoliata and systemic infection in Nicotiana benthamiana"		
	join(108..7997,7999..9381)		
	/note="The polyprotein encompasses domains for papain-like leader proteinase, putative methyltransferase, RNA helicase and RNA-dependent RNA polymerase (RdRp); ORF1a/b; papain-like leader proteinase is capable of autoproteolytic release from the rest of the polyprotein, and is a replicational enhancer. RdRp is presumably expressed via a +1 ribosomal frameshift"		
	/codon_start=1		
	/product="p348"		
	/protein_id="AAFI4300.1"		
	/db_xref="GI:6492368"		
	/translation="MAFLNLSAVPSCAFAPAFAPAGASPTVPDSFFPCVPRYSDDISHRLTSLDQSVPRPLFNLNHLRASTNPLSLPLGHFACTFVLELNGSSAPSPISRLHDFVNRPSPEPTEVLVSSTDSRLCLDFEYCSKQPCVLEIASFTSPPLCLVNSVQAQIETHAMESIRPFTYLPAGRLQEHKKYTKRPETLIHESGLAKVSPFFGTSPKNSQPTTVKASGSKYEAYEISLRKDFERSRRQRTPRVSHKPKINKAVEHAPVAVSGKVRTPFRTAGETVTAAGYVNRPAQTSRPNHPIILRSASYTFGEKMPVKKHMEKQYVVKRSKVYSSCSVTSPLEALTSLKLNPRYSNRSRLKFDHPHFDGDFEIT		

```

EVHLRGRGKLSVLLIPXGEAYCVVTAATPQVHAACTIARGDRPRVGVGLLQYRPGEG-
CYLAHAALCCA-OKRTFEEEDPFGVGYTKFVAKRLTEKLGPSALXHPVAGROVSRS
LFHCDVASFPSPFYSJPRFGVVEEBAPEITSS:KHAIAIESVVERVS:HKDNLARS
VEKOLDFPKDE:KSLKSKRSVTVFFNGEAVQSGLTRAYFCFNLSPFTHSVSDHPAA
AGSLJEWETUASAKSSFDIGGCPPLPHIKRGSDTHVCRPIYDMKDAQRVSRE:Q
ARGLVENISREQLVEAQRVSCPTLGNKGVKSDVLIVGVQYDASNEIASAMVJKE
SKVAYLTVTPELLEDEREAIDALGCVVDTRDMQVQKFGPSSCYCHKLNIKMI
MLTAPFTSGNLFSEVMENMGVNYKI:TRSAYSPEIRGKVTLYRACACTEVSVCVL
PREDKTLFDGXDYIYLDKAFYRSRVFVYVUSNCVSVNSKTFEWMWSYIKSSKRVV
ITSGKVIHRODHLKHSCEFAVLAUGVRSTT:FEIAKNLNITYDASCFETIRFL
FREWGRYAEINRFRMLKMSIISAGUDYIFLDDNSLOHLJEYSVEVRSIAQNG
VEVCKENRVSJTE:IAEADAKRSIAQ:SGALSVPTQRGJRGGRSRSVSLLYNL
VEEONLFTGLRLDIFENFLKSKLVIAENACFAYIERNMKLYFPPLNSBEEAR
EAVSFAFSWADRTVESGVKTVNRTVKRFLNFMASKTLTKKFRFF:ASALAKTVV
RKAKVILAEYMEVWPESITLSDSGEYSAYFCSSVITLITNSGRLLPGFSFSAVITV
LLDLATKISIEVLLKQISFVUSTASSALYRRVLSEILLSNFRMGEHGIFIKVFLUCGF
LPVVRKVALUCVPGDMATYARFLEYGVDDIFFLGRSVNSIKNYLCVVAAGVDS:VD
SPVJLKGAVKERVLGFKSKIKNFLNVRKAKVITRTSSSTDLSEDEYFSCDSKPG
LRGSSRFTLRLDIFENFLKSKLVIAENACFAYIERNMKLYFPPLNSBEEAR
RLIRICAGDFDY:SDSAPDEDMROAFQYSSDDSEVTDGKPTVLRSYLVNSRRFL
ETFCGPKPFVKNVYFKALYSRLVLVWDRNLSDSPGLKGGNEKALLAKFLKTCV
ITACEVSOICCLRLILRCWGPAGGLVRLFYITYSGTRVLRSVVUVAVCPLLVRNE
LDGSDGTLNMGVSFVRKLFVALRRALSAYSALRRKIFEFIFGNTHHPEDVAVIE
NEVAPPELPEVDIDVDCDSDSESSEVASIPRPLRGGRSRSNFLTSLVKVY
FKLAGRIPLRLFRUNFAYIVERRLASKR:KTFIGLARLDFNLSLTVVYLLQEYDS
V:NAFIDVELVLNAGSVNVLPLVSWVRGSLTKLAEVIVGSGFASFLGRMCCRVSQDW
SSSNAGCNFPVPR:KGPVPPSSGTSAGMVELEALEDIREHVLSTCRVSGDE
EERPEKVEPGIETSEDVPLRSHSOPLSGGECSSYEDREENBRANLLPHVSKIVSE
PRGLETABRNKRTLRGUSEELNAINTSNEQRPPIIVDHSPPSRALTSREFYVJQBL
ALFELSCKLRYIOLQKLVANFRNCECLDKOEDFVLKAGGVSGRNSRULPHFKG
HEFCFRSGGLVPYGTSVDVTFHTQTFNVSANALLSGYLSYRTFTFNLSANVJJE
APPGKGTTLIKVFCEFTKVNLSILITANKSREEILAKVNRIVJDEGDTPLQTRR
ILITDSVLNMRGJTKVLYLDECFMVGAGAAVACIEFTKCDSAILFQDSRLQHYDR
NELDTAVISDLNRFDVDESVRVGEVSYRCPWQVCAM:STFVPTKVATNLNVSAGQSM
QVRETESVDVDEYSEFVPLWLOSEKDLKSEFGKRSRSSVEKPTVLTVEAGCEY
KRYNLVR:KQEDDPRFRENHITVALSRHVESYYSVSSRRDRAICAAIKVAIKVJY
AIRVPTSPGSDTIDISWNPSTDRSKKASSAPYEVINSFLESVPFGTTSVDFGDYS
SEMG:QVPSGACNVVIRDSAPVNSKTDHPQVRSISRSQAIPKRKFSLENLYSES
RNYNFTVCEPFGOEFGQAMAMVLERFDEKVAKYRSDVI:AITKGVRTWMSKRE
PSQALRSLQKLOKPLEEETITFKLMVKRDAKVLDSCLVKLPHPPAQNIMEHKAVN
AIFSPCDEKRVITCTNSINIFTEWNTSLASIAKEM:GSEHVYVNSGELDEKED
KSQDAFIKSFERTLYSAGFPEDDLDMVQGEYTSNATLLOGLSFSDVQKSGASN
TWIGNSIVTGLNSFTYINRAFKLVSGGDSLIFSPSPRNSADACTEGEFTKFL
TPSVYQCKFTMGHTYFVDPQKLVKLGASKDEVDDEFLFVEFTFRDLTKL
VDERVIELTHLHSXYGYESDYYAA:CAIHC:R3NF:SSFPKLYPKVKGWVHYGKL
KFLVRKPFANCFREKFD:AFGEAYFLTYDET"
9445..9609
/note="small hydrophobic protein; ORF2"
/codon_start=1
/product="p6"
/protein_id="AAFI4301.1"
/db_xref="GI:6492369"
/translation="MDCVLRSVLLAFGLFLICFLFCVVFIMFYKQLFRNTPPSN
EARNRSTVV"
3609..11405
/note="HSP70 protein homolog; ORF3; required for
cell-to-cell movement"
/codon_start=1
/product="HSP70h"
/protein_id="AAFI4302.1"
/db_xref="GI:6492370"
/translation="MVVFGDLFGTTFSSVCAYVGEELYLFKQDASVYPTFVFLHSPT
QEVAFGYDAEVLNDSPVRGFGYRDLKRWIGCDDEENGYDLEK:KPKYKTELLKVAS
SKSTVRLDCYSOTVPQNAITPLGLATFAVKALISTASEAFQCQCTGVCSIPVANCYQ
RSFTESCYNLGSYPCVVMVNPSPSAALSACSRIGKATSPVLAVDFGCGTIDNVISAL
NNTFVVRASGGDMGLGSDIDKAFVHELYKKAQLPNYK:DISFLKSLSKSVFLNF
PVWSEQNVKDVUVNSELAEVAPVETIKIVKEVTEKRGSRMRLEFSPVAKLLKY
GSSV:LPGLLRSLSSVPFVECLVLPDARAAGVGCALYSACLNRDPMFLVADCAAHN
LSISSYCEISVCIVPAGSPIPTGVTINVMAGSNASVYSAALFEGGFVCKRLNRF
SGDVALGVGVGSGNTRVPLTLEINVSSTVGTITFLVGPVGVKKLVGNGAAVDFSSY
QLGERVWADLHKHNSDKVKLIHALTYKPEQRKKT:SDSKALFLKRLSADYREAGKFS
SYDDAVTNSSELGRLORVPIKILGRSGREKLV"
11311..12972

```

```

/note="64 kDa protein; ORF4"
/codon_start=1
/product="p64"
/protein_id="AAFI4303.1"
/db_xref="GI:6492371"
/translation="MTTRSTPANYNGELFRFFFGGRENKLN:MSFAASVRRPVTDRD
FRFSNGEVRKNSFSDSTGESFVRFSFLLLPKTYEVCKL:GVAMEQALSQMNRLSD
YVNSLNVVDVTKTVCKENIQTUTVFRKLNGN:AEPSIVHECMSLNSNSCGELIKPKD
TRFVSLIFPKDQVVESTDEAVVSSYLOYLSHCLNLYETCNLSNGKALYDFLK
YVIVLETSDLETRSUSDNPLVAGVLYDMCFYNT:KSTLYKNIESFCFSLYDPLJ
SEIFSMWQOPADPVLSDLELDTTELLIPTINTHSDTFLYKNIKLY:LESYFEDDSN
EITKVKVDSLLTRDNPDLKLAQRVFGHCYGVFRTAQTRKVKRDAEYKLPALGEFT
INMSGVEEFDELQKMPSPSVRRPCGSL:SHAEPISFKRPGVGPPI:TRLNVPKYS
YLVNDVYRVRKAGLTQDELITLSNIEFDVAEMCCEREVALOARQORKEK:FPQWKG
VKNESYRRASISIRVKSNSLNLVNLKVDVGARRQGRNLPLRKH"
12920..13570
/note="ORF5; CPM"
/codon_start=1
/product="minor capsid protein"
/protein_id="AAFI4304.1"
/db_xref="GI:6492372"
/translation="MLAPEGRGLIHPTENTRDAMET:FNSYDIAEYSEVNKNLNRK
ETDELGVIRERFKSELVITDEDFVKHLAFALIRANITTSKVANYGAYEYTGKK
FLVKDAWVPLJ:KECKKFNKPNVTFCACTEDAYIVIARSLPKLFLNRTIKRGIP
SGYFEGADPL:ATSVCLNDHEKAIVLQASRAAIDRAVSSSVGDKIVSFLGRSL"
13641..14255
/note="ORF6; Cp"
/codon_start=1
/product="major capsid protein"
/protein_id="AAFI4305.1"
/db_xref="GI:6492373"
/translation="MGSABPISATFENVSLADQTC:HLHGECDCDKLRRDFECLKLG
VPEDK:GLAGLGLCLYSCATIG:SNKVSVOPTSTPIKASFGSGKELFTHGELRSFLDS
OKLLEKPNKLRCPQRTFOKDYISPAKEYRGLPLPIARNRRLPAEDHY:LAADPST
STELTDQGRLLLRARENTHTEFSSESPTSLKQLGRGLATGK"
14252..14785
/note="20 kDa protein; ORF7"
/codon_start=1
/product="p20"
/protein_id="AAFI4306.1"
/db_xref="GI:6492374"
/translation="MTSSVELLAQTRPLFRVLLKGFVYIVAIETEESEPEVELPLVY
LHDLNVAKKGIKESSYIDFXSCMTRLXPSVSYTRVSSGXPSDFSSSYSGKTFDS
KILNRKVTTFESGILQVFGMYGRDQRCVSSEYLVFENFVFGAHCGTLPYCLNCELDK
SGOEILFTSKNEVJ"
14769..15302
/function="enhancer of RNA accumulation"
/note="21 kDa protein; ORF8"
/codon_start=1
/product="p21"
/protein_id="AAFI4307.1"
/db_xref="GI:6492375"
/translation="MKKEFENDCETRA:SRSESL:PRVKELCGNSQSSQSEVSEKINEF
ELARFNHLVTVHEHREMEXHPKQSELRTPRJGEMLKEIRAFKRVVITPMHETA
SETLNALFEEYCRITGLTRDALRERKRVSTVLFHSELLKFEVTNMFSTELLK
LNLRLVSISS:LGIAV"
BASE COUNT 3866 a 347: c 3697 g 4434 t
ORIGIN

Alignment Scores:
Pred. No.: 6,05e-21 Length: 15468
Score: 297.00 Matches: 71
Percent Similarity: 54.50% Conservative: 32
Best Local Similarity: 37.57% Mismatches: 84
Query Match: 29.97% Indels: 2
D5: 14 Gaps: 2

US-09-613-486-15 (1-198) x AFI90581 (1-15468)
QY 7 Ser:AsnLeuSer:Asn:LeuValIleThr:AspAlaSer:SerLeu:AsnGlyVal:AspLysLys 26
Db 13671 GCGACTTTTGAACACCGTGAGTCTTCGACGACCAACCGTGTTCGACCGTGAAGACTCGGAC 13730
11311..12972

```

```

QY      27  LeuLeuSerAlaGluValGluLysMetLeuValGluLysGlyAlaProAsnGluGlyIle 46
DB      13731 AAATCGGAGGATTTGGAAGAGTGTGTTGAAATGAAAGGGGTTCCGGAAGATAAATCTC 13790
QY      47  GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAargThrThrSerProLysVal 66
DB      13791 GGCTCGCGGTAGGACTTTGTTGTACTCTGCTGCGCAGCATAGTACTTCTAATAAGTT 13850
QY      67  GluArgAlaAspSerAspValIlePheSerAsnSerPheGlyGlu---ArgAsnValVal 85
DB      13851 AGTGTCACACCGAGCTCTACTTTTCATCAAAAGCTTCGTTCGGTAGTGGGAAGGAATTTCTC 13910
QY      86  ValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105
DB      13911 CTCACTACAGGTGAAGTCTCTTCTGAGTCTTCAGAAAGGATTACATATCTCTCGCAAGGAA 13970
QY      106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPheCysIleAla 125
DB      13971 AACAAATTGGGTGTTTCTCGCGCACTTTTCAGAAAGGATTACATATCTCTCGCAAGGAA 14030
QY      126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
DB      14031 TACCGAGGAAGATGCGTCCGATTCGTAGAGCAACCGTCACCGTCTACTCTGCTGAAGAT 14090
QY      146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnSer 165
DB      14091 CACTACTAGTCTGATTTTCATATCGACATCACAGAACTTACTGACCTACACAGGCT 14150
QY      166 ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrPro 185
DB      14151 CGTCTGCTGTGGCGCGGAAACCGCCATCACACAGAAATCTCTCG---TCTGAATCACCA 14207
QY      186 ValSerAsnLeuArgGlnLeuGlyArg 194
DB      14208 GTAACAGTTTGAACAGCTGGGTCTG 14234

RESULT 7
LOCUS      BYU71295
DEFINITION Beet yellows virus coat protein gene, partial cds.
ACCESSION  U71295
VERSION     U71295.1
KEYWORDS    GI:1619939
SOURCE      Beet yellows virus
ORGANISM    Beet yellows virus
REFERENCE   1  (bases 1 to 612)
AUTHORS     Creamer,R., He,X.H., Yang,C.H. and Grantham,G.
TITLES      Characterization of the 3'-proximal encoded proteins of beet yellows closterovirus
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 612)
AUTHORS     Creamer,R., He,X.H., Yang,C.H. and Grantham,G.
TITLES      Direct Submission
JOURNAL     Submitted (19-SEP-1996) Plant Pathology, University of California, Riverside, CA 92521, USA
FEATURES    Location/Qualifiers
             source          1..612
                        /organism="Beet yellows virus"
                        /mol_type="genomic RNA"
                        /isolate="California BYV-CA"
                        /db_xref="taxon:12161"
             CDS             1..>612
                        /codon_start=1
                        /product="coat protein"
                        /protein_id="AA017001.1"
                        /db_xref="GI:1619940"
                        /translation="MSSAEPISAIATFNENVLYVDOTCLHGECDCDKLRNFEELCKLKG
VPEDKGLA.GLCLYSCATIGTGNKYSVQPTSTFFKASFGSGKELFTHGEIARSLFDS
QKLEGGPNKLCFRFTQKDYISFAKEYRGLRFPPIARNRNHLGPAEDHYLAADFIST
STELTDQQGRLRLARENATHTEFSSDSPVTSIKQLGRGLATGD"
BASE COUNT 165 a 144 c 147 g 156 t

```

```

ORIGIN
Alignment Scores:
Pred. No.:      1,09e-22      Length:      612
Score:          296.00      Matches:    71
Percent Similarity: 54.50%      Conservative: 32
Best Local Similarity: 37.57%      Mismatches: 84
Query Match:    29.87%      Indels:      2
DB:             14          Gaps:        2

US-09-613-486-15 (1-198) x BYU71295 (1-612)

QY      7  SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26
DB      31  GCGACTTTTGAACCGTGAAGTCTCTGAGACCAACAGTGTTCGACGGTGAAGATCGGAC 90
QY      27  LeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIle 46
DB      91  AAACCTACGAGGAATTTTCGAAGAGTGTGTTGAAATTTGAAAGGGGTTCCGGAAGATAAATCTC 150
QY      47  GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAargThrThrSerProLysVal 66
DB      151  GGTCTCGCGTTAGGACTTGTGTTACTCTCTGTGCGACGATAGGTACTTCTAATAAAGTT 210
QY      67  GluArgAlaAspSerAspValIlePheSerAsnSerPheGlyGlu---ArgAsnValVal 85
DB      211  ACGTCCCAACCGAGCTCTACTTTTCATCAAAAGCTTCGTTCCGTAGTGGGAAGGAATTTGTTTC 270
QY      96  ValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105
DB      271  CTCACCTACGCGCAACTGAGGTCTCTTTCTGCACTCTCAGAAACTTTTGTAGAGGGAAGGCT 330
QY      106  AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPheCysIleAla 125
DB      331  AACAAATTGGCGTGTTCGCGCACATTTTCAGAGGATTACATATCTCTCGCAAGGAA 390
QY      126  TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
DB      391  TACCGAGGAAGATGCGTCCGATTCGTAGAGCAACCGTCACCGTCTACTCTGCTGAAGAT 450
QY      146  SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnSer 165
DB      451  CACTACTTAGCTGCTGATTTTCATATCGACATCACAGAACTTACCGACCTACCAAGGT 510
QY      166  ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrPro 185
DB      511  CGTCTGCTGTGGCGCGGAAACCGCCACTCACAGAAATTTCTCG---TCTGATTACCA 567
QY      186  ValSerAsnLeuArgGlnLeuGlyArg 194
DB      568  GTAACAGTTTGAACAGCTGGGTCTG 594

RESULT 8
LOCUS      AF056575
DEFINITION Beet yellows virus California isolate, complete genome.
ACCESSION  AF056575
VERSION     AF056575.1
KEYWORDS    GI:3283076
SOURCE      Beet yellows virus
ORGANISM    Beet yellows virus
REFERENCE   1  (bases 1 to 15468)
AUTHORS     Peremyslov,V.V., Hagiwara,Y. and Dolja,V.V.
TITLES      Genes required for replication of the 15.5-kilobase RNA genome of a plant closterovirus
JOURNAL     J. Virol. 72 (7), 5870-5876 (1998)
MEDLINE     98285746
PubMed      9821048
REFERENCE   2  (bases 1 to 15468)
AUTHORS     Dolja,V., Peremyslov,V. and Hagiwara,Y.
TITLES      Direct Submission

```

```
JOURNAL Submitted (31-MAR-1998) Botany and Plant Pathology, Oregon State
University, 2082 Cordley Hall, Corvallis, OR 97331, USA
FEATURES
Location/Qualifiers
1..115488
/organism="Beet yellows virus"
/mol_type="genomic RNA"
/isolate="Californian"
/db_xref="taxon:12161"
-08..9381
Gene
/gene="ORF1a/b"
CDS
join(108..7983,8045..9381);
/gene="ORF1a/b"
/codon_start=1
/product="fusion protein"
/protein_id="AAC25115.1"
/db_xref="GI:3283077"
/translation="VAFILNVSAVSCAPAPAPAPAHAGASPIVPSFPCVPKRYSDISH
FRILUSLDFSVPRPLFNARVHLRASTONPUSPLPLGHASTFVLELNGSSAPFESIPS
RHIDFVNRPSPFTEVLSVSSLRTPSLFALLCDFLYGSKPGCPVEIASFSTPPP
CLVSNCAQIPTHAMESIRFETKLPAGRLQLQHKRYTKRPETLI1HESGLAKTS
ALGVTSKPNRPITVKSASGKYEAYEISRKDFSRRRQOTPRVRSHKPKIKNAVE
PFFPEEPKDKRKRASLPTKDEGPIFTGLRFLPSETPKPEPRLPKPREVEIPVKK
HAPRAVSKPVPTFRPVATTGAEVYNAPTQCSRPRNHPILRSAYITGFKMPLQRP
MKEKDYVYKSKSVSSVTKSPLEALTSILKMLPRYSINSEKLFYDHF1GDOFEI
SVPLRUGKLSVILLIPKGEAYCVVTAATPQHAALT1IARGDRPRVGGELQYRPGEL
CYLHAALCCALQKRTFEEQFVGMPYTKFVFAKRLTEKLGPSALKHPVRGRQVRS
LPHCDVASAFSPFYSLPRFFIGVBEAEPIITSSLKHAIESVVERVSI1HKDNLARS
VEKD1IDFDEIKSLSKKRSVTPVPMGEAVOSGLTRAYPQFNLSFTHSVYSDHPAA
AGSRILENETIASMAKSFSFDIGGCP1PHIKRGSTDVHVRP1VYMDAOPRVSRELO
ARGLVEN1SRELOEACARVSCVHTLGNCKVKSVDL1MQVQYDASLNEIASANWLE
SKYALVT1PQSLJDEERA1DALGCDVDVDRVQYKFGSSG1CHXLSKIKNI
MLTPAFTSGN1FVSEMYENRGVNYK1ITRSAYSPSEIRGVKTLRYRAACTEVKQVL
PRFDLTKTFLSGYDIY1LDAKFSRVDFVYVNSGVNSK1TFEWMASY1KSSSRVY
1SGVK1IHRDVIHDIKHSFCFAAMLAGVGRSRTT1FEFLAKNLVYTGASCFETIRPL
FEMWCKRAEY1NRRGFKMLKMS1LSAGLGYEFLDNLNQLHLEYSEVEVRVSAQNG
VEDNCEENRVLTET1AARADRS1JAQGLSGLSVPC1QPRGLGGRSRSVGSFLYD
VEEVGNL1FESGVDAVRF1LVKVPKTSFSDSPI1FRVRMLD1JAAAS1FVSVSLCAWLR
EAVAFSSWADRTVSSSVKTPVNR1TKRFLNFMASKTLTKK1FEFFFLSASALAKTV
RAKAV1LEAYWEVFEFS1LSDSGESAYEFCSSVVI1TLTNSGLLPGFSESATIEY
LIDLATKIS1EYLLQI1SPVSTASSALRYRVL1LNSFRMTGHEGIFTKVF1LCGF
LPVFRKVCVACPGDMATYAR1FLEGYDD1FLGRSVNS1KNYLCVVAAGLVDI1VD
SVVLKSVAKER1ELVFGKSK11KNF1NVPRKAVTVTRTS1SD1SEDEVEFSCDESKPG
LRGSSRRT1SEL1DLFENELKSK1V1ENACFSAYE1ERNMKLY1FPLNSSEBAR
RLIRACAGDF1YSDAFDEDEMLRQAFQSVSDSESV1YDGKPTVLRSLY1NLSRRE1
ETFCGVSQ1FVKVSNY1FALYSLRLLRVL1PWDRN1JSDSPG1KGGNEKAL1AKF1KTCV
ITAC1CVSQ1CCLRL1RCWTPACGLVRL1FYITY1SGTRV1SRVYVAVAV1CPL1VRNE
LDG1SDPLT1NMGVS1FRALF1VALR1AAL1AGNSAL1R1FEFI1GN1HHPDVAIET
NEVAP1EL1PEVD1DVCDFGSDSESVSDSEVAS1PRFLG1GGGSRSSN1PLTS1VKVY
FK1ACR11RLL1PRLPNVAV1VERE1ASKRL1KTR1GLAF1FCNPSL1TSVWV1LQEYS
VLNA1T1DELV1JNSGSVNV1PLVSVKRGSL1TKAEV1VSGFAS1FGRXC1CRVSDWC
SSS1NAGCNFMS1PVTKK1FVPPSSGSTASKYER1LEAESDI1REHVL1LSTCRVSDSE
EER1KXE1TEPGIET1SEDV1RSHSQ1PLSGECS1YSEDREERAN1L1PHVSK1VSE
RGLT1ABRNKR1T1RGVSEFLNAIN1SNEQRP1I1VDSHSPESRAL1TSVREFFYLQEL
ALFELSK1RLRY1DOLKVAN1EOR1QCLDCKDE1MFLRAG1GVVSGNRSL1PLKHXD
HEFC1FRSGGLVPYDCT1SDVT1FHTQTN1FVSANAL1SGY1SYRTFTFN1SANKVLYE
APPGGK1TTL1KVCET1FSKVN1SL1TANKSREEL1AKNVR1VLDEGDT1PJQTRR
LITD1SY1LNNRGL1CKVLY1DECFVSVHAGA1VAC1EFTKCD1SAIL1FGDSQ1HIY1DR
NELO1TAV1LDRNFV1DDESV1VGYRCPMD1CWLST1FYPKTVAT1N1VSAQSSM
QVRE1SDVD1VESEF1VLTMLQEK1DLK1SKFGKRS1RSVEXPT1VLT1VHAQOETV
RKVN1VR1TK1FOBD1P1FRSENH1TVALSRH1VESLTV1SVLSSKRD1DAI1ACAI1VKAKOLVD
AVRVY1PT1SGSLD1L1SVNPST1SDRSK1KASAP1VEVINS1FLES1VWPGT1SVDFG1DVS
EEMGT1QVPSG1NDNV1R1D1SAPVNS1KSTHEEN1LYS1ESRNY1NFTV1CER1FSG1QFQGA
MANW1KERS1F1DEK1VAK1RVS1AIETK1GVRT1MMS1KREPS1QRLAS1SDLP1KPLNEE
1ITFF1TMTN1STFLAG1IAKEM1LGSHEV1N1VGE1IDFS1KDK1SQDAP1KSP1EFLYS1AGFG
DEDL1D1VW1QGG1TSNAT1LDQ1LS1FSD1NOR1KSGAS1N1TGN1SIV1TL1GILSM1EYV1TN
RFLK1FV1SGD1SL1SES1P1RNS1AD1XCT1ELG1FETK1LTP1SVY1FCS1K1F1WMT1CHV1F
FVPD1PY1LL1VKLG1ASK1D1VDE1F1FEV1T1F1RDL1DK1LD1VIEL1LTH1V1SK1GYE
SGDT1AAL1CA1H1C1R1NS1FSS1FKK1Y1PKV1GW1V1VY1GK1FL1VLR1K1FAN1C1FRE1K1FAT1F3
EAY1FLTYDET"
108..1871
/gene="ORF1a/b"
/product="papain-like leader peptide"
```

CDS  
QKLEGPKNLRCFCRTFKDYISFAKEYRGLRPPIARANRHHGLPAEDHYLAADFIST  
STELTDLQOGRLLARENAHTTFSSSPVTS.KQLGRGLATCK"  
14252..14785  
/function="unknown"  
/notes="ORF 7"  
/codon\_start=1  
/product="20 kDa protein"  
/protein\_id="AAC25121.1"  
/db\_xref="GI:3283083"  
/translation="MSSVELAQTKPLFRVLLKGFVFVVALETBEESPEVELPLVY  
LHDLNKNKGIKSIYDFKSMTRLKSPSSVSYRVSSGKPSDEFSSYSGKTFDS  
KILNRKVTFTFSGICLVFGMGYRQDCVSSEYLFNFVFGAHGCTGYCLNCELCX  
SGSELEITPSKNEVEL"  
14769..15302  
/function="replication enhancer"  
/notes="ORF 8"  
/codon\_start=1  
/product="21 kDa protein"  
/protein\_id="AAC25122.1"  
/db\_xref="GI:3283084"  
/translation="MKFFPNDCTSRASRSSELRLRVKELGTNSPQSEVSECINEFN  
ELARPNHLVTVVHRERMEKHPQSSLELTPSELGMLKEIRAFJAVRVVPMHETA  
SETLNAFLSEYCYRTGLTRDALREMRKRVSTVLFHSELKFEVTENKFSFTELLK  
LNLSJRVISSQILGIAY"

BASE COUNT 3874 a 3472 c 3696 g 4426 t

## Alignment Scores:

Pred. No.:	9,93e-21	Length:	15468
Score:	295.00	Matches:	71
Percent Similarity:	53.97%	Conservative:	31
Best Local Similarity:	37.57%	Mismatches:	85
Query Match:	29.77%	Indels:	2
DB:	14	Gaps:	2

US-09-613-486-15 (1-198) x AF056575 (1-15468)

Qy	7	SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys	26
Db	13671	CGCACTTTTGAAACGCGAGTCTCGTAGACCAACACGTTTTCACGCGTAGACTCGCAC	13730
Qy	27	LeuLeuSerAlaGluValGluLysMetJeuValGlnLysGlyIleProAsnGluGlyIle	46
Db	13731	AAACTACCGAGGAATTTTGAAGAGTGTGTAATTTGAAGAGGGTTCCGGAAGATAAATC	13790
Qy	47	GluValValPheGlyLeuLeuLeuTyThrAlaLeuAlaIleAlaIleThrSerProLysVal	66
Db	13791	GGTCTCGCGTAGGACCTTGTGTTGCTCTCTGTCGACGATAGTACTTCTAATAAAGTT	13850
Qy	67	GlnArgAlaAspSerAspValIlePheSerAsnSerPheGlyGlu---ArgAsnValVal	85
Db	13851	AGCGTCCACCGACGCTACTTTTCATCAAGCTTCGTCGTCGTCGTCGTCGTCGTCGTC	13910
Qy	86	ValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaProLeuThrArgPheThr	105
Db	13911	CTCACTCACGCGAAGTACGATCGCTCTCTGACTCTCAGAAATCTTTTAGAGGAAAGCCT	13970
Qy	106	AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyThrValAspPheCysIleAla	125
Db	13971	AACAAATTCGGTGTTCGCGGACCTTTTCAGAGGATTAATATCTCTCGCGAAGGAA	14030
Qy	126	TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp	145
Db	14031	TACCGAGGAGACTCGCTCCGATTCAGAGCAACCGTCACGCTACCTGCTGCTGAAGAT	14090
Qy	146	SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer	165
Db	14091	CACACTAGCTGCTGATTTTCATATCGACATCAACAGACACTTACTGACCTACACACAGG	14150
Qy	166	ArgLysMetPheAlaSerMetTyThrAlaLeuLysThrGluGlyValValAsnThrPro	185
Db	14151	CGTCTGCTGTTGGCGCGAAGAACGCCACTCACACAGAAATTCGCG---TCTGAATCACCA	14207

Qy	186	ValSerAsnLeuArgGlnLeuGlyArg	194
Db	14208	GTAACCTAGTTTGAACACAGCTGGTCTG	14234

RESULT 9  
BYVCPG 2724 bp ss-RNA linear VRL 02-AUG-1993  
LOCUS Beet yellows virus coat protein gene, complete cds, and two ORFs.  
DEFINITION M59452.1 GI:323237  
ACCESSION M59452.1  
VERSION coat protein.  
KEYWORDS Beet yellows virus  
SOURCE Beet yellows virus  
ORGANISM Beet yellows virus  
VIRUSES: ssRNA positive-strand viruses, no DNA stage;  
Closteroviridae: Closterovirus.  
REFERENCE 1 (bases 1 to 2724)  
AUTHORS Brunstedt,J., Moseley,J. and Hull,R.  
TITLE Nucleotide sequence of cDNA encoding the coat protein of beet yellows virus  
JOURNAL Unpublished (1990)  
COMMENT Original source text: Beet yellows virus cDNA to genomic RNA.  
FEATURES  
source  
1..2724  
/organism="Beet yellows virus"  
/mol\_type="genomic RNA"  
/isolate="Broom's barn"  
/db\_xref="taxon:12161"  
CDS  
1..1287  
/codon\_start=1  
/label=ORF\_1  
/protein\_id="AAA72953.1"  
/db\_xref="GI:323238"  
/translation="PDFVRKLNGLAEPSLVHCHMSNSCGELINPKDKRFPVSLIF  
KKGWVESTDAVYSSVLDYLSHCLNLYETCNSSNGKKALYDEPLKYVIAVLET  
DLEVRSPDNELVAGLYDMCFEYNTLSTYLNKLESPDCFLSLYLPJLSIFSMNWE  
QPADVRLPEEDTAELSLKIPTINTLSTFYLNKRDABEYKLPALGETIINMSGVEEF  
LTDNPELKLQARWVGPHCYGVFRFTAQTRKVRDAEYKLPALGETIINMSGVEEF  
FDEKQKPVSVRRFCGSLSEAFSIFKRGVGFPPITRLNVPVKYSYLNVDYR  
VRAGLTODELTILSNIBFVAEMCCEREALQARRAQKPKFGQWKGKNEVSPHA  
RSSIRVKSNSLNLWKDVGARRQRLNPLRRKH"  
1235..1885  
/codon\_start=1  
/label=ORF\_2  
/protein\_id="AAA72954.1"  
/db\_xref="GI:323239"  
/translation="WLAPEGRDLTHFTENTRDAMETFNSSVDLAEYSEVNPKNLRK  
ETDELGVIRBRKSELVITDEDFVKHLAFALIRAGN:TTTKNYVUAYETIGKK  
FLVKDAWFFLIKCKMKRPNPNVPTFCATFEDATVIARSLPKLNLRNRTIGKGIP  
SGVEFLGADFLTATSVCLNDHEKAIVLOASRAAIDRAVSSSDVGKIVSLFDLGRLS"  
1956..2570  
/codon\_start=1  
/label=ORF\_3  
/product="coat protein"  
/protein\_id="AAA72955.1"  
/db\_xref="GI:323240"  
/translation="MGSAPETSAITAFENVSLADQTLHGDCCKLRKNFECLKLG  
VKDELKGLALGLCATIGTSNKSQVPTSTFIKASFGGKELFTHGLRSLDS  
OPKLEKGNKLRFCORTFKQDYISFAKEYRGLRPPIARANRHHGLPAEDHYLAADFIST  
STELTDLQOGRLLARENAHTTFSSSPVTS.KQLGRGLATGR"

BASE COUNT 739 a 591 c 638 g 756 t  
ORIGIN

## Alignment Scores:

Pred. No.:	1.64e-21	Length:	2724
Score:	293.00	Matches:	71
Percent Similarity:	53.97%	Conservative:	31
Best Local Similarity:	37.57%	Mismatches:	85
Query Match:	29.57%	Indels:	2
DB:	14	Gaps:	2

US-09-613-486-15 (1-198) x BYVCPG (1-2724)

Qy 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26

Db	1986	CGGACTTTGAAACGTCAGTCTCGCAGACCAACAGTGTTCGACGCTGAGACTCGGAC	2045
QY	27	LeuLeuSerAlaGluValGluGlyMetLeuValGlnGlyysGlyAlaProAsnGluGlyIle	46
Db	2046	AAACTACGGAAGAAATTCGAAAGAGTGTTCGAAATGAAAGGGTTCGGAAGACAAATC	2105
QY	47	GluValValPheGlyLeuLeuTyrAlaLeuAlaAlaArgThrThrSerProLysVal	66
Db	2106	GGTCTCGCGTTAGACATTTGTTGTATTCCTGTCGACGATAGTACTTCTAATAAGTT	2165
QY	67	GlnArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal	85
Db	2166	AGTGTCCCAACCGACGTCTACTTTTCATCAAAAGCTTCGTCGTCGGAAGGAATGTTC	2225
QY	86	ValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThr	105
Db	2226	CTCACTCAGGTGACCTAGAGTCTTTCTTGACCTCAGAACTTTTGGAGGAAGCCCT	2285
QY	106	AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPheCysIleAla	125
Db	2286	AACAAAGTGCCTGTTCTGCGGCACITTTTCAGAAAGCACTACATATCTTCGCGAAGAA	2345
QY	126	TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp	145
Db	2346	TACCGAAGAACTGCCTCCATTGCTAGAGCCCAACCGTCACCGTCTACCTGCTGAAGAT	2405
QY	146	SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer	165
Db	2406	CACACTTAGCTGCTGATTCATATCGACATCAACAGAACTTACTGACCTACCAACAGGT	2465
QY	166	ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrPro	185
Db	2466	CGTCTGCTGTTGCGCGCGGAAAGCCCACTCACACAGAGTTCG---TCTGAATCACCA	2522
QY	186	ValSerAsnLeuArgGlnLeuGlyArg	194
Db	2523	GTAAGTATTGAAACACGCTGGGTCGT	2549
RESULT 10			
BYVMBPA			
LOCUS	BYVMBPA		
DEFINITION	Beet yellow virus (isolate German BYV-3) genes for putative	linear	VL 01-OCT-1993
	membrane-binding protein, heat shock 70-related protein, coat		
	protein homolog and coat protein.		
ACCESSION	X73475.1	GI:313689	
VERSION			
KEYWORDS	coat protein; heat shock 70-related protein; membrane-binding		
	protein.		
SOURCE	Beet yellow virus		
ORGANISM	Beet yellow virus		
	Viruses; ssRNA positive-strand viruses, no DNA stage;		
	Closteroviridae; Closterovirus.		
REFERENCE	1		
AUTHORS	Agaravsky,A.A., Koonin,E.Y., Boyko,V.P., Maiss,E., Lunina,N.A. and		
	Atabekov,J.G.		
TITLE	Beet yellow closterovirus: complete genome structure and		
	identification of a leader papain-like thiol protease		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 5980)		
AUTHORS	Maiss,E.		
JOURNAL	Direct Submission		
	Submitted (18-JUN-1993): E. Maiss, Inst. f. Biochemie und		
	Pflanzenvirologie, Biologische Bundesanstalt, Institut fuer		
	Biochemie, Messweg 11/12 D-3300 Braunschweig, FRG		
FEATURES	Location/Qualifiers		
source	1. .5980		
	/organism="Beet yellow virus"		
	/motype="germic RNA"		
	/isolate="German (BYV0G)"		
	/db_xref="taxon:12161"		
	/clone="pBYV-D22, pBYV-D31"		
	71. .235		
CDS			

	/note="putative membrane-binding protein"	
	/codon_start=1	
	/protein_id="CAA51857.1"	
	/db_xref="GI:313690"	
	/db_xref="SPTREMBL:O65893"	
	/translation="MDCVLRSYLLAFGLICLFLCLVVFIFWVYKQILFRNPPSN	
	EARENRSTVV"	
	235. .2031	
	/codon_start=1	
	/product="heat shock 70-related protein"	
	/protein_id="CAA51858.1"	
	/db_xref="GI:313691"	
	/db_xref="SPTREMBL:O65894"	
	/translation="WVVG-DRGTTFSVCAYVGEELYFKORDSAVPTFVFLHSPT	
	QVAFGYDAVLSNDSSVRGFTYRDLRNIGCDSEENVGYDLKLRHYKTELLKVAQS	
	SKSTVKDCTSGVTPQATILPGLIATFVKALISTASEAFKQCTGVTCSPVANYCQ	
	RSFTSCVNLSGYPCVTVMNEPFAAALASGRKIGATSPVLVYDFGGTIDVSVISAL	
	NNTFVRAAGDMMNGRDIDKAAPVHLRYKQLPVNYKIDISFLKESLSKVYFNLN	
	PVSEQNVKDVILVNSAEVAAPFVERTIKIVNEVYKGRMSLEPVLKQALMV	
	GGSYLQGLLSRSSVPFEELVLPDARAAGGACALVACLRNDSPMILVCAAHN	
	LS-SSKYCESTVCVPAGSPITFTGVTVMNMGSNASAVYSAALFEGDFVKRLNKR1?	
	SGDVALGDVGFGNSNTRTVFLTLEINVSSTGTTTFLVGPTVKKLVGGNAADFFSY	
	QUGSRVADLHKNSDKVLIHALTYKPFORKKLTDSKALFLKRLSADYRRARKEFS	
	SYDDAVLNSSELLGRVIPKILRGRVEKLDV"	
	1937. .3598	
	/note="orf4"	
	/codon_start=1	
	/protein_id="CAA51859.1"	
	/db_xref="GI:313692"	
	/db_xref="SPTREMBL:O65895"	
	/translation="MTTFTSTPANYMGELFRFFGGREWNLMSEASVRRFVYARD	
	FRFSGVLRKIPFSDTGESFVREFSLILTPFKTYEVCKLCGVAMEQALSGMNRISD	
	YNSFELNVAVDKTVGCKFNQTVTEFVKLGNLAEPV-VEHCWSLSNSCGKALINPKD	
	TKRSLIIFKGDIESTDBAIVSSVLDVLSHCLNLAYETCNLSNSGSKALVDEFLK	
	VYIAYLSDLEYSRSDNBLVAGLVDMCFEYNTLKYTLKNIESPCDCLSLYLPJ	
	SEIFSNWBPAPDRLFLDLTAELLLKPTINTLSTLYKNKRYLESIFEDDSN	
	ELIKVQSLSLTRDNPDLKLAQWVGPHCYTGFRTAQTRKVRKDAEYKPCPALGET	
	INMSVEEFDELQCKMFSVRRRFGCSLSEAFSIFRFGVGPPIITLNVPPVRY	
	YLNVDYHYVRKAGLTQDELITLSNIEFDVAEMCCEREVALQARRAQRREKPPQGNKG	
	VKNEVSPHARSIRVKKSNESLLNLWKDVGVRROGRNLNLRKH"	
	3546. .4196	
	/codon_start=1	
	/product="coat protein homolog"	
	/protein_id="CAA51860.1"	
	/db_xref="GI:313693"	
	/db_xref="SPTREMBL:O65896"	
	/translation="MLASEGRGLIHPTENTRCAMETFFNSYDLAEYSEVNPKNLNRK	
	STDELLGVRERPKSELVITDEDFVKHLAFLIRAAKITSTKVNVYGAVEYIGKKK	
	FLVADNVFPLIKCKMKKKNPRTFCATFEDAYIVARSIPKFLARTIGKRIE	
	SGYEFGLADRLTATSVCLNDHEKAIVLQAGRAAIDRAVSSSDVKIVSLFDLGRLS"	
	4257. .4881	
	/codon_start=1	
	/product="coat protein"	
	/protein_id="CAA51861.1"	
	/db_xref="GI:313694"	
	/db_xref="SPTREMBL:O03724"	
	/translation="MGAEEPSAIAIFENVSLADQTCJHGDCDCKLRNFEBCLKJUG	
	VPEDLGLALGLCYSCATIGTSNKVSQPTSFIRKASFGGKELFTTHELESFLDS	
	QKLEGPKNLRGCFRTFKQDYISFAKEYGRPLFIARANRHGLPAEDHYLAADFIST	
	STELTDLQOGRLLARENA-HTBFSSESPTSLKQJGRGLATOR"	
	4878. .5411	
	/note="orf7"	
	/codon_start=1	
	/protein_id="CAA51862.1"	
	/db_xref="GI:313695"	
	/db_xref="SPTREMBL:O65897"	
	/translation="MIGSVELAQTRPPFRVLLKGFVFIYVIAIETEEPESEAEPLPVY	
	LHDELNVNKKGIESYIDFKSCMRRLKPSVSIVRVSGKPSSEFSSYSKSTFES	
	RILNRKVITFENGTLQVPGMYGRDRRCVSSEYLVWFENFVGAHCGTLFYCLNCELDK	
	SGDLEILTFISKNEVFL"	
	5395. .5928	
	/note="orf8"	
	/codon_start=1	
CDS		

/protein\_id="CAA51863.1"  
/db\_xref="GI:313696"  
/db\_xref="SPTREMBL:Q65938"  
/translation="MKFFNDGESRAISBSLLRVKELGNSQSEVSEGINBPN  
ELAFNHLVVERHEWKEHPKQSSURAPSRJGEMLKEIRAPLKVVRVVPYKHSTA  
SETLNALDEYCRJAGTREDALREKMRKRVSTVLFHSELLKFEVTENNFSFTLJK  
LNLSRVISSQIGIAV"  
BASE COUNT 1557 a 1267 c 1431 g 1725 t  
ORIGIN

Alignment Scores:  
Pred. NO.: 4,64e-21 Length: 5980  
Score: 293.00 Matches: 71  
Percent Similarity: 53.97% Conservative: 31  
Best Local Similarity: 37.57% Mismatches: 85  
Query Match: 29.57% Indels: 2  
DB: 14 Gaps: 2

US-09-613-486-15 (1-198) x BYVMBPA (2-5980)

QY	7	SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys	26
DB	4297	CGGACTTTGMAACGGTGGCTCCGAGACCAACGGTGTTCACGCGTGAAGACTGCGAC	4356
QY	27	LeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIle	46
DB	4357	AAACTCGGAAGAAATTCGAAGAGTGTAAATTTGAAGGGGTCCGGAAGACAACTC	4416
QY	47	GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaIleArgThrThrSerProLysVal	66
DB	4417	GGTCTCGGTAGACACTTTGTTGCTATCTCTGCGCAGTAGACTACTTCTAATCAAGTT	4476
QY	67	GlnArgAlaAspSerAspValIlePheSerAsnSerPhe--GlyGluAsnValVal	85
DB	4477	AGTCGCCAACCGAGCTGCTACTTTTCATCAAGCTTCGTCGGTGGGAAGGATGTTTC	4536
QY	86	ValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaProLeuThrArgPheThr	105
DB	4537	CTCACTCAGCGTGAACCTGAGTCTTTCTGGACTCTCAGAACTTTTGGGAAGGAACCT	4596
QY	106	AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPheCysIleAla	125
DB	4597	AACAAGTTGGTGTGTTCTGCGGACCTTTTCAGGAAGGACTACATATCTTCGCAAGAA	4656
QY	126	TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp	145
DB	4657	TACGAGAAAGACTGCTCGGATCTTAGAGCCACCGTCACGCTCACCCTGTGAGAT	4716
QY	146	SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer	165
DB	4717	CACCTACTAGCTGCTGATTCATATCGACATCAACAGAACTTACTGACCTACAACAGGT	4776
QY	166	ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGlyValValAsnThrPro	185
DB	4777	CGTCTGCTGTGGCGCGCAAAACGCCACTCACACAGAGTTCTCG---TCTGAATCACCA	4833
QY	186	ValSerAsnLeuArgGlnLeuGlyArg	194
DB	4834	GTAACCTAGTTGAACAGCTGGGTCTG	4860

RESULT 11  
CLBYV3PH 6746 bp RNA linear VRL 13-JUN-2002  
LOCUS  
DEFINITION Beet yellows virus genome 3'-proximal half, for capsid protein  
HSP70 related protein, RNA-dependent RNA polymerase and ORFs.  
ACCESSION X33462  
VERSION X33462.1  
KEYWORDS GI:58878  
capsid protein; heat shock protein 70 homologue; RNA polymerase;  
RNA-dependent RNA polymerase; unidentified reading frame.  
SOURCE Beet yellows virus  
ORGANISM Beet yellows virus  
Viruses; ssRNA positive-strand viruses, no DNA stage;  
Closteroviridae; Closterovirus.

REFERENCE	1	(bases 1 to 6746)
AUTHORS	Agranovsky,A.A., Boyko,V.P., Karasev,A.V., Lunina,N.A., Koonin,E.V. and Dolja,V.V.	
TITLE	Nucleotide sequence of the 3'-terminal half of beet yellows closterovirus RNA genome: unique arrangement of eight virus genes	
JOURNAL	J. Gen. Virol. 72 (Pt 1), 15-23 (1991)	
MEDLINE	91116305	
PUBMED	1990061	
REFERENCE	2	(bases 1 to 6746)
AUTHORS	Agranovsky,A.A., Boyko,V.P., Karasev,A.V., Koonin,E.V. and Dolja,V.V.	
TITLE	Putative 65 kDa protein of beet yellows closterovirus is a homologue of HSP70 heat shock proteins	
JOURNAL	J. Mol. Biol. 217 (4), 603-610 (1991)	
MEDLINE	91171285	
PUBMED	2005613	
REFERENCE	3	(bases 1 to 6746)
AUTHORS	Rensing,S.A. and Mater,U.G.	
TITLE	Phylogenetic analysis of the stress-70 protein family	
JOURNAL	J. Mol. Evol. 39 (1), 80-86 (1994)	
MEDLINE	94343547	
PUBMED	7545947	
REFERENCE	4	(bases 1 to 2677)
AUTHORS	Agranovsky,A.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (22-MAY-1990) Agranovsky A.A., A.N. Belozersky Laboratory, Moscow State University, Moscow 119899, USSR	
REFERENCE	5	(bases 2678 to 6746)
AUTHORS	Agranovsky,A.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (27-JUL-1990) Agranovsky A.A., A.N. Belozersky Laboratory, Moscow State University, Moscow 119899, USSR	
FEATURES	Location/Qualifiers	
source	1..6746	
	/organism="Beet yellows virus"	
	/mol_type="genomic RNA"	
	/db_xref="taxon:12161"	
CDS	1..659	
	/codon_start=3	
	/product="RNA polymerase"	
	/protein_id="CAA37549.1"	
	/db_xref="GI:808947"	
	/db_xref="SPTREMBL:Q66109"	
	/translation="DNQRKSGASNTWIGNSITETLGLSKFYVYTRFKALFVSGDDSLI PSESPIRSADAMCTELGFEKFLTPSPYPCSPFVMTGHDVFPVPYKLLVKLGA SKVDDEPFEVFTSPFDLTDLDERVIELTHLVHSKYGESGDTIYALCAIHCI RNSFSSFKLJPKVKGWVHYGKGLFVLKFCANCFREKEDTAFGERTFLTTLKLETVL "	
CDS	710..874	
	/note="ORF2"	
	/codon_start=1	
	/protein_id="CAA37550.1"	
	/db_xref="GI:58879"	
	/db_xref="SPTREMBL:Q08542"	
	/translation="MDCVLSYLLAFGLICLFCLVFWFVYKQILFRTTAQSN EARHNHSTV"	
CDS	874..2670	
	/codon_start=1	
	/product="HSP70 homologue"	
	/protein_id="CAA37551.1"	
	/db_xref="GI:58880"	
	/db_xref="SWISS-PROT:P37092"	
	/translation="MVVFLDFTGTFSSCAVYGEELYLFKQDSAYIPTVYFLHSDT QEVAFQDAEVLNDLSVRGFGYRDLKRWICDDENRDYLKUKPHYKTBLLKVAQS SKSTVKLDCYSGTVPQNATLPLGIATFKALISTASEAFKCCQCTGVCISPAVNCLO RSFTESCYNLGYPCVYMNPEPSAALSAACSRKIGATSPVLVDFGGTGFVSVISAL NNTFVRASGDMLGGRIDKAFVHLYNKAQLPVNYKIDISFKESLSKKVSLNF PVSSQGVYVDVLVNVSELAFAAFVETIKIVKVVKYCSSMRLPVPYKAKLIMV GGSYLPGLLSRLSIPFDVDECLVLPDARAAGGALYSACLNDSPMLLVDCAHN LSISKCESIVCPVAGSPVFTGRTVTNMTGSNASAVYSAALEFEGFVFKCRKLRIF FGDVVLGNVGTGATRTVLTLEINSSVGTISFLVGTGFKKLGGNAAYDFSSY QLGERVADLHKNSDKVLIHALTYQPPQRKLTGDKALFLKRLADYRREARKFS SYDDAVLNSSELLGRITIPKILGRSVEKLDV"	

CDS

```

2576..4237
/note="ORF4"
/codon_start=1
/protein_id="CAA37552.1"
/db_xref="GI:58881"
/db_xref="SPTREMBL:Q66110"
/translation="MTTRESTPANYWYMGELFRFRFGQEWKMLSEAAVSRRPVSSD
FRFSDGVILSRKTFGSGTGVSRFESLTLTPKTYEVCKLCVAMEALANGWELSD
YVSEFNIYDVMTVGKFNQVTPFKVINGNAEPLVCHMSLSKSELINPKD
TRPVSLIPKGDJLAEDEBIVSSYLYLSHCLNLETJCNSSNGKSLIDFJK
HVIDLENSDJLRSPSDNPLVAGILYDMCFEYNTLSTYLNIESFCFFSLYDPL
SEFSNWERPADPVRLLFELDAEALLKVPFTINMHDSTFLYKRLRYLESFEDDSN
ELIKVVDLSLLTRDNELKLARWGFHYGVYVTRTAQTKVRDAEIKLPLGEFV
INMGVEEPEELCKXKMPISVRRPCGSLSEAFSVKFGVGPPTIRLNVPKYS
YLNVDYRVRKVGILGQDELTLISNLEFPAEMCCEREVALOARQRRPFCQKKG
TKNEISPHARSSIRVKNDLSLLNLIKVDVGARSQRRLNPLRKH"
4185..4835
/note="ORF5"
/codon_start=1
/protein_id="CAA37553.1"
/db_xref="GI:58882"
/db_xref="SPTREMBL:Q68538"
/translation="MAAPARGDLIHFTENTRDAMETPFNSVDLLEYSEVRPNKLNK
ETDELGVIREKPSLVITDEDFVKHLAPALIRANITTSVKVYGVAYETIGGK
FLVKDAMVPLIKCKMKXKNPNVPTFCATFEDAYIVTARSPLKFLNRTIGRGIP
SGVEFGASPLTATSVCLNDHEKAIVLQASRAIDRAVSSVDGKIVSLFDLGRLS"
4906..5520
/codon_start=1
/product="capsid protein"
/protein_id="CAA37554.1"
/db_xref="GI:58883"
/db_xref="SPTREMBL:Q68531"
/translation="MGSAEPIGAIATFENNVSADOTCLHGDCCKLKNFEBCLMLKG
VPEDNLGIALGLCYSCATIGTSKNVVOPTSTFIKASFGGKELYLTHGELNSFLGS
KXLECKPKNLRFCGTROKDIATSLRKEVKGKLPETARANRHCPRADHYLAADIST
STESTDQQRJLLARENA:HTFEFSSESPTSLKQLGRGJGTGR"
5517..6059
/note="ORF7"
/codon_start=1
/protein_id="CAA37555.1"
/db_xref="GI:58884"
/db_xref="SPTREMBL:Q68544"
/translation="MTSSVELAQTKPLFRVLLKGFVTVVAFETEESSSEAEPLVY
LHDFELNKRKGIKIASYVFMSCVTRJRPSSVSVTRVSSSEDFSLPGSGKTFGS
KVLNRKVTTFENGVLQVFGMYGLEQRCVSSDYWFENFVFGAHGCTLYCLNCEJDK
SGGELEILTFKNEVLLKRW"
6034..6567
/note="ORF8"
/codon_start=1
/protein_id="CAA37556.1"
/db_xref="GI:58885"
/db_xref="SPTREMBL:Q68545"
/translation="YKFFLDGDETSALRSSESLLRKVRKELGTNSQOOSRISCVDFEN
ELASFNEJLVTVHREHMEQHPNOSKLRVPSRIGMLKRIEAFILKRVVITPWEKTA
SDTLNAPLEYCRITGLAREDALREKVRKSVLPHHSELKFEVTENMFSYTBLLK
LNLISRVISQILGMAI"

```

BASE COUNT 1775 a 1415 c 1581 g 1975 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3 08e-20 length: 6746  
Score: 286.00 Matches: 71  
Percent Similarity: 52.91% Conservative: 29  
Best Local Similarity: 37.57% Mismatches: 87  
Query Match: 28.86% Indels: 2  
DB: 14 Gaps: 2

US-09-613-486-15 (1-198) x CLBV3PH (1-6746)

Qy 7 SerAsnLeuSerAsnLeuVali.eThrAspAlaSerSerLeuAsnGlyValAspLysLys 26  
Db 4936 GCGACTTTTCAAAACGTAAGTCTCGCAGACCAACCACTGTTGCGACGGAAGACTCGCAT 4995

```

Qy 27 LeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGlyLe 46
Db 4996 AACCTTAGGAGAACTTCGAGAGAGTGTGAAATTTAAAGAGGGTTCGGAAGATAACCTC 5055
Qy 47 GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAargThrThrSerProLysVal 66
Db 5056 GGAATCGGTAGGACCTTGTGTTGATTCCTGCTACGATAGGCACCTTCCAAACAAAGTT 5115
Qy 67 GlrArgAlaSerAspValIlePheSerAsnSerPhe---glyGluA:ArgValVal 85
Db 5116 AACGTCCAAACCGAGCTACCTTCATCAAAAGCTTCGTTGCTGTCGGAAGAACTGTAC 5175
Qy 86 ValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThr 165
Db 5176 CTCACTCAGCGTGAATTCGTTCTGCGGTCTCAAAAACCTTTTGAGAGGAAAACCT 5235
Qy 106 AsnLysLeuAargThrPheGlyA:ArgThrPheThrGluAlaTyrValAspPheCysIleAla 125
Db 5236 AACCAATTCGGGTGTTTCTGCGGTACTTT-CAGAAAGGACTACATATCCTTTCGCGAAGAA 5295
Qy 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
Db 5296 TACCGAGGGAATTTACCTCCGATTGCCAGAGCTAACCGTCACCGTCTACCCGCTGAAGAT 5355
Qy 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
Db 5356 CACTACTTAGCGCTGACTTCATATCGACGTCGACGAACTCACTGACCTACACAAAGT 5415
Qy 166 ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrPro 185
Db 5416 CGTCTGCTGTAGCGCGCAAAACGCCACTCACACGGAATTCCTCG---TCTGAATCACCG 5472
Qy 186 ValSerAsnLeuArgGlnLeuGlyValG 194
Db 5473 GTACCAAGTTGAAACACACTAGGTCGT 5499
RESULT 12
LOCUS BYVJAA 15480 bp RNA linear VRL 19-MAY-1995
DEFINITION Beet yellows virus (isolate Ukrainian BYV-U) complete genome.
ACCESSION X73476
VERSION X73476.1 G:405624
KEYWORDS coat protein; heat shock 70-related protein; heat shock 90-related protein; methyltransferase; papain-like protease; RNA dependent RNA polymerase; RNA helicase.
SOURCE Beet yellows virus
ORGANISM Beet yellows virus
Viruses; ssRNA positive-strand viruses, nc DNA stage;
Closteroviridae; Closterovirus.
REFERENCE 1
AUTHORS Agratovsky, A., Koonin, E.V., Boyko, V.P., Maiss, E., Frotscher, R.,
Lurina, N.A. and Atabekov, J.G.
TITLE Beet yellows closterovirus: complete genome structure and
identification of a leader papain-like thiol protease
JOURNAL Virology 198 (1), 311-324 (1994)
NEEDLINE 94082464
PUBMED 8259666
REFERENCE 2 (bases 1 to 15480)
AUTHORS Agratovsky, A.A.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1993) A.A. Agratovsky, A.N. Belozersky Institute,
Moscow State University, 119899 Moscow, Russia
COMMENT Related sequence: X53462.
FEATURES
Location/Qualifiers
1..15480
/organism="Beet yellows virus"
/moltype="genomic RNA"
/isolate="Ukrainian (BYV-U)"
/db_xref="taxon:12161"
/cclone="pBYV p36, x19, R9, R3, 36A, R2, 36, p43, 111-7,
112, 115, 1213, 1210, 124, 1313, 1311, 142, 143, 1311,
1311, 142, 143, 156, 1510, 154, 1518, 1520, 169, 1615, 175"
1..107
5'UTR

```



gene	join(108..7983,8045..9393) /gene="ORF1a/b" join(108..7983,8045..9393) /gene="ORF1a/b" /codon_start=1 /product="fusion protein of papin-like protease, methyltransferase, RNA helicase" /protein_id="CAA51871.1" /db_xref="GI:8035339" /db_xref="SP:TRMBL:Q08534" /translation="MAFLNVSVPSCAFAPAFAPAHAGASP:VPDSFPCCVPKRYSDDIASH FRULSLDFSVPRLSNARVHLASTDNPRLSLPLGFAETFVLELNGSSAPFSPS RHIDFVNVNPFVFTPELVSLNRTSPSLFALLCDFLYCKPGPCGVKSTFPP CLVSNCAOIP2THAEMESIRFPTKLPAGRFQFHKKYKRPETLIIHESGALKTS ALGVTSKPNRSP:TPVSNASGEKYEALSRKOFERSRQQTPTVRSHXPRINKAVE PFFPEPKDKAKRASLTEDGGF:FGTLRFPUSETPKPEPRJPKFREVEIPVVKK MAKVAVYKRVKSVSSGVNARNOCRPNRPHILRSASYFGFKMPLCRF EYHPLRGKLSVLLIPKGEAYCWTAATPQYHAALTIARGDRPRVGEHLQYRPGEG CYLHAALCCALQKRTFREDFVGVNYPKVFHAKRLTEKLGSALKHPRVGRCLARS LFHCQVASAFSSPYSLPRFQIGVEEAEPTTSSLKHKALESVYGERVSTHKNLARS VENDLIDPFDEHSLSKERSVTVPYMGEAVQSGLTRAYPOFNLSITSHVSDHPAA AGRLLENELASMAKSSFD:GGCFPHIKRGSTDIHVCRPIYDCKAQRVSRRELQ ARGLVNLREQUVEAQAQAVCPHTLGNVKSVDLIMVQYDASLNEASAMWLKE SKVAYLTVTPGELLDEREFAFADALGCVVJTRDMVQYKFGSGCYCHKLSIKSI MLTPATFTFNGNLFSDMYENRKNVYKITSAYSPETRGVKTLYRACVYKSGSRVY PRFDKLTFLSGDYVYLDAQKFSRVFDVYVSNCSVSNKTFEMWASYIKSKSRVY ISGVLIHROVHIDLKHSECFAAVMJAVGVRSTTTFSLAKNLYWTDASCRET:RFL FWRSSRAEAINRSFKLAKMSILSAGLDYFELDLNLSQHLSEYSEVRVYSIFANG EVDCEENRVLTIIEAADRKSLIAQGLSALSSVPTQPGRLGSRSSRVSFLYNL VEEVLNLFSGDAVFLVKVFETSDSPIFRVRFNFDJLAEAAFPFVSVSLCAWLK EAVSAFSSWADRTSVESKVTFNRTVRFPLFMSAKTLTKKFFRFLSASLAKTV RKAQVILEAVMEVWFESILSDSGEYSAVFCSSVWILTNSGRLLPGFSPSAILTEV LLSLATKISIEVLKILSDPADSTASSALYRVLSEILSNFRTWGHEGCTFKYVFLCGP LPVFKVCALCVFGGMATYRFLLEYVDGDFELGSRVNSIKNYLCVAAAGLVDSIVC SWLKSUGVAKERVLGFKSKIIKNFLNFRKAKVTRTSSDLSDEYFPLNSDEEAR LRGSSRFTSLRLDIFNFKLSKJVIENACFSAYER:ERNMKLYFPLNSDEEAR RIIRCAGDPFYSDADFDEMLRQAFQVYSDSDSVYCGKPTV:RSLVNVSRFL ETFCNGKPFYKSVNFKALYRLLRVLFWDRNLSDSPLGKGEKALLAKFFKTV ITACEVSOIOLCCRLRLCWGT PACGLVLEVTYSSTRVLRVVAVAVACLLVRYNE LOG:SGDLNMGVSVFRRLVALRRLASVNSALRKKIET:FGNIHPPVDVAVIET NEVAPLESPEVDIDCDPFGSDSESVDVANSPRGLGSGRRSSNFL:SLVKVY FKLARIPELLFLRNFVAYFVERRLASRKLKTFI GLARLDFNFTSVYLLQYDS VLFNAFIDVELILNCSGVNVLPLVSVWRGSLTKLAEAVGSGFASFGMCCRVSQDW SSSNAGCNFMSPVRTKGAFPVPSSSGSTASMYERLEALESDIRBHVLSCTQSGDEE BERPKVTEGIEHTSEDVVP:IRSHSQPLSGGECYSDEERENRANLILPHYSKIVSE RGLETARNKRTLHGUSFELKNAINTSNQPRFIIVDHSPESKALTN:SVRETY:QEL ALFELSKUREYIDQLKVFANFRQECIDKQEDMFVIRAGQGVSGRNSRLPLKHPKG HEFCFRSGGLPYDGTSDVTIFHTQTNFVSNALLSGYLSYRTFTNLSANLLYE APPGGKTTLLIKVFCETFSKVNLSLITANKSREBILAKVNRILVDEGDTLQTRDR ILTIDSYLNMNRLCTCKVLYLDCEFMVHAGAAVACIEFKCDASAILFGDSROIYGRCL SELCTAVLSDIARFVDSRVSRYRCPMVD/CAMLSFEPKVTATNLVSAGQSSM QVREISVDVDEYSFEVYLTMLQSEKKDLVLSFGKRSSEVEKPTVLTVLAQAGETY RKVNLVTRKQEDDPFRSENHITVALSRVLESUTSVLSKKDDAIAIVAKAKQVLD AVRYPTSGGTLVDVSVNPS:SDRSCKKASAPAEVINSFLESVWPTSTVDFGDSV EEXTQVFGSGADNVYIRDSAPYKSTDEHENLYSYESRNVNFTVCRFSGSQBEQQA MAMVLEBSRDKLKVAKRSDVIAITEKGVRTMWSKREP:SQLRALSDDQKPNLLEE IITFKLMVKRDAKVLDDSCLVKHPPAQMIPHRKAVNAIPSPCFDEPKRVITCTNS NIVVFTMNTSLIASAKMGLSEHYNVGEIDFSKEDKSQDAFIKSFERTLYSAGFP DELLDVWGCETYSNATLJGGCLFSFSDNORKSGASN:WIGNSITLGLSK:FYITN RFKALPQSGDLSIIFESPRIANSADAMCTELGETFJTPSPYFCSKFFV:GHGVF FVDPYKLLVJLKGASDEVDDELFEVFTFSFRDLTXD:VDERV:ELLT:HLVHSKYGYE SGTGYAALCAHICIRNFSSFKLYPKVKGWVYHYGKLFVLRKFPANC:RENKFTAFG ERTFLTLTKLETVL"	exon	8045..9393 /gene="ORF1a/b" number=2 9444..9608 /note="putative membrane-binding protein" /codon_start=1 /protein_id="CAA51866.1" /db_xref="GI:405627" /db_xref="SP:TRMBL:Q08542" /translation="MDCVLSYLLAFGLICLFLCLFLCJLVWFVFWYKQILFRTT:AQSN EARNHNSHTVV" 9608..11404 /codon_start=1 /product="heat shock 70-related protein" /protein_id="CAA51865.1" /db_xref="GI:405628" /db_xref="SWISS-PROT:P37092" /translation="MVVFGLDFTGTFSSVCAYVGEELYLFKQRDSAY:PTVVFUHSDT QEVAFGDAEVLNDLSVRGQFYRLKRWIGCDENVRDYLEKLPKPHYKTLLKVAQS SKSTVKLDCYSTVPQNATLPGLTATFKVALISTASEAFKQCQCTGVLCSPVAVNCLQ RSPTESCUNLSGYPGVVYMWEPSPAALSACSRIKGATSPULVYDFGSGTFDVSIAL NNTFVSAGSDGMLNGRDIDKAFVEHLYNKAQUPVNYKIDISFLKESLKKVSLFNP PVSEQGVRVLDLVNVSLEAEVAPFVERTIK:VKVEYKCYSSMRUEP:PNYKAKLMMV GGSSYFGLLSRLSSIIPFDECLVLPDARAAGGCALYSACLRRDSEFMLVDCAAHN LSISKCYCESIVCVPAGSP:IPFTGVRTVNMFTGSNASAVYSACLFEQGFVKRLKRIIF FGDVVLGVNVTGSATRTVPLTLEINVSSVTI:SFSLVGP:GVKKL:GGNAAVDFSSY QLGSRVADLGHKNSDKVLIHALTYQPRPKKLTGDGKALFLKELIADYRREARKFS SYDAVALNSELLILGRILPKILGRSVEKLDV" 11310..12971 /note="putative heat shock 90-related protein" /codon_start=1 /protein_id="CAA51866.1" /db_xref="GI:405629" /db_xref="SP:TRMBL:Q08544" /translation="MTTRFSTPANYWGEUFRFRFGQEWKUNLSEASVSRRPYSSD FRYSGVILSRKTEGSTGESFVREFSLLTTPKTYECKVLGAMELANGMRLND PNFSDFINVDVKTGVGCKTQSVTEFVKKINGVNAEPLSLVHCWMSLNSCEELNPKD TKRFSV::FKGDLAESTDEAIVSSYLDYLSHCLNI:VETCNLSNGKGLSYDEFLK HVIVDLENSDLEYSPSDNPLVAGILYDMCFYNTLAKTYLKNIESDCCFLSLVPLL SEVFSMWERPAPDVRLLLFELDAELLKVPYTIMHDSSTELYKNKLRYLESYFEDDSN ELIKVDSLUTRONPELKAQRWVGHCYVGFRTAQTRKVKRDAKYKPPALGEFV INMSGEFFFEELQKMP:SVRRFCGSLSHSEAFVFRFGVGFPPIITRLNVPKYS LVNDYVHVRVRLVGLTQDELITLSNIFDVAEMCCEREVALQARRAQERGEKPFQWKG TKNYSIPHARSIRVKKNN:DSLLAILMKDVGARSQRRLNPLRKH" 12919..13569 /codon_start=1 /product="coat protein homolog" /protein_id="CAA51867.1" /db_xref="GI:405630" /db_xref="SP:TRMBL:Q08538" /translation="MLAPEARGDLIHENTRDMAMETFFNSYDLAEYSEVNPKNLNRK ETDELLGVIRBRPKSELVITDEDPVKHLAFALIAANITTSVKVNYVGAET:YIGGKK FLVMDVPLIIECKMAKFNPNVPTFCATDEDAIVIAKSLPLFLNRTIGKRGIP SGYFLGADFLTATSVCLNDHEKAIVLQASPAADRAVSSVDGKI:VSLFDLGRJS" 13640..14254 /codon_start=1 /product="coat protein" /protein_id="CAA51868.1" /db_xref="GI:405631" /db_xref="SP:TRMBL:Q08531" /translation="MGSAPISAIATFENVSLADQTLCHGEDCDKLRNFESECLKLG VPELNLGIALGLCYSCATIGTSNKNVQPTSTIKASFGGGKELY:THGLNLSFLGS QKLEGNKRLCFRTFOKDYISLRKEYRQGLPPIARANRHGLPAEDHYLAADFIST STELTDLQQSRELLARENATHTEFSSSPVTS:LQGLRGLTGR" 14251..14793 /note="orf7" /codon_start=1 /protein_id="CAA51869.1" /db_xref="GI:405632" /db_xref="SP:TRMBL:Q08544" /translation="MTSSVELAQTKPLFRVLLLKGFVYVAVFETETEESSEAEPLVY LHDFELINRKGKITEASGYDFMFSOMTRLKFSVSYSYTRVSSEKSSDLPQSGKTFGS KVLNRKVTFTFENGVQLVFGMYGLEQRCVSDYLWFENFVGAHCGTLTYCLNCELDK	CDS		misc_signal	7983..8045 /gene="ORF1a/b" /note="+1 ribosomal frameshift signal" 7984..8044 /gene="ORF1a/b" number=1	intron	
------	---	------	--	-----	--	-------------	--	--------	--

```

CDS
SGSELEILTSKNEVLLRW"
14768..15301
/locus="orf8"
/codon_start=1
/protein_id="CAA51870.1"
/db_xref="GI:405633"
/db_xref="SPTREMBL:Q08545"
/translation="MKFFADGTSRALSRSLELLRRVKELGTNSQCSISCVDFN
ELASFNLLVTVHEHREMEHQNSKLRVPSRIGEMLEIRAF.LKRVTVTPMKETA
SDTLNATFEYCRITGLAREDALREKMKVKSVLFHHSLLKFEVTENMFYSITELLK
LNLRLRVISQILGAI"
15299..15480
3'UTR
BASE COUNT 3891 a 3446 c 3680 g 4463 t
ORIGIN
Alignment Scores:
Pred. No.: 9,23e-20 Length: 15480
Score: 286.00 Matches: 71
Percent Similarity: 52.91% Conservative: 29
Best Local Similarity: 37.57% Mismatches: 87
Query Match: 28.56% Indels: 2
DB: 14 Gaps: 2
US-09-613-486-15 (1-198) x BYVCAA (1-15480)
QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26
Db 13670 CGGACTTTGAAACGTAAGTCTCGCAGACCAACCTGTTTCACGAGGAAGACTCGGAT 13729
QY 27 LeuLeuSerAlaGluValGluMetLeuValGlnLysGlyAlaProAsnGluGlyIle 46
Db 13730 AAACCTTAGGAAGAACTTCGAAGAGTGTTCGAAATTAAGGGGTTCCGGAAGATAACCTC 13789
QY 47 GluValValPheGlyLeuLeuTyTAlaLeuAlaLaargThrThrSerProLysVal 66
Db 13790 GGAATCGCGTTAGACACTTTGTTGATATCCGTGCTACGTAGCACCTTCCAAACAAAGTT 13849
QY 67 GlnArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal 85
Db 13850 AACGTCCACCGACGCTACCTTCATCAAGCTTCGTTGCTGGTGGGAAGAACTGTAC 13909
QY 86 ValThrGluGlyAspLeuLysValIleAspGlyCysAlaProLeuThrArgPheThr 105
Db 13910 CTCACCTCAGCGTGAATGAAATCCCTTCTCGGGTCTCAAAACCTTTTGGAGGGAAACCT 13969
QY 106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyValAspPheCysIleAla 125
Db 13970 AACAAATTCGGTGTTTCTCGCGTACTTTTCAGAAGGACTACATATCTTCGCGAAGAA 14029
QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
Db 14030 TACCAGGGAATTAACCTCCGATTCGCCAGAGCTAACCGTCAACCGTCTACCCGCTGAAGAT 14089
QY 146 SerTyLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
Db 14090 CACTACTTACCGCGCTGACTTCATATCCAGCGTCGCGAGACTCACTGACCTACCAACAGT 14149
QY 166 ArgLysMetPheAlaSerMetTyTAlaLeuLysThrGluGlyGlyValValAsnThrPro 185
Db 14150 CCTCTGCTGTGTAGCGCGGGAACGCCACTCACACGGAATTCCTG---TCTGAATCACCG 14206
QY 166 ValSerAsnLeuArgGlnLeuGlyArg 194
Db 14207 GTAACAGTTTGAAACAACTAGTCTGT 14233
RESULT 13
A41914
LOCUS A41914 783 bp DNA linear PAT 05-MAR-1997
DEFINITION Sequence 1 from Patent WO942847.
ACCESSION A41914
VERSION A41914.1 GI:2297469
KEYWORDS Beet yellows virus
SOURCE Beet yellows virus
ORGANISM Beet yellows virus
Viruses; ssRNA positive-strand viruses, no DNA stage;
Closteroviridae; Closterovirus.
REFERENCE
1 (Bases 1 to 783)
AUTHORS Bojzen.K. and Brunstedt.J.
TITLE DNA SEQUENCE COMPRISING AT LEAST TWO COAT PROTEIN GENES
JOURNAL Patent: WO 9428147-A 1 08-DEC-1994;
SANDOZ AG (AT)
COMMENT OTHER Publication AU 7122694 941220.
FEATURES
Location/Qualifiers
source
1..783
/organism="Beet yellows virus"
/mol_type="genomic DNA"
/db_xref="taxon:12161"
BASE COUNT 223 a 176 c 171 g 211 t
ORIGIN
Alignment Scores:
Pred. No.: 2.3e-21 Length: 783
Score: 285.00 Matches: 70
Percent Similarity: 53.44% Conservative: 31
Best Local Similarity: 37.04% Mismatches: 86
Query Match: 28.76% Indels: 2
DB: 6 Gaps: 2
US-09-613-486-15 (1-198) x A41914 (1-783)
QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26
Db 119 CGGACTTTGAAACGTAAGTCTCGCAGACCAACCTGTTTCACGAGGAAGACTCGGAC 178
QY 27 LeuLeuSerAlaGluValGluMetLeuValGlnLysGlyAlaProAsnGluGlyIle 46
Db 179 AAACCTAGGAAGAAATTCGAAGAGTGTTCGAAATTCGAAGGGTTCGCGAAGAAACATC 238
QY 47 GluValValPheGlyLeuLeuTyTAlaLeuAlaLaargThrThrSerProLysVal 66
Db 239 GGTCTCGCGTTAGAACCTTTGTTGATATCCGTGCTGCGACGATAGGTACTTCTTAATAAGTT 298
QY 67 GlnArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal 85
Db 299 AGTCTCCACCGACGCTACTTTCATCAAGCTTCGTTGCTGGTGGGAAGAAATGTTC 358
QY 86 ValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105
Db 359 CTCACCTCAGCGTGAACCTGAGGTCTTTCTGACTCTCAGAAACTTTTGGAGGAAAGCCT 418
QY 106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyValAspPheCysIleAla 125
Db 419 AACAAATTCGGTGTTTCTCGCGCACTTTTCAGAAGGACTACATATCTTCGCGAAGAA 478
QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
Db 479 TACCAGGGAAGACTGCTCCGATTCGTAGAGCCCAACCGTCAACCGTCTACTCTGCTGAAGAT 538
QY 146 SerTyLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
Db 539 CACTACTTACCGCGCTGATTCATATCCAGACTCACAGAACTTACTGACCTACACAGGT 598
QY 166 ArgLysMetPheAlaSerMetTyTAlaLeuLysThrGluGlyGlyValValAsnThrPro 185
Db 599 GGTCTGCTGTGTAGCGCGGGAACGCCACTCACACAGAGTTCTCG---TCTGAATCACCA 655
QY 186 ValSerAsnLeuArgGlnLeuGlyArg 194
Db 656 TAACTAGTTTGAAACAACTAGTCTGT 682
RESULT 14
BYU51931
LOCUS BYU51931 10545 bp RNA linear VRL 05-APR-1999
DEFINITION Beet yellow stunt virus helicase gene, partial cds; and RNA
replicase, p30, p6, p86, p61, p25, coat protein, p18, and p22
genes, complete cds.

```

J51931 L20761  
 US1931.1 GI:1388128  
 Beet yellow stunt virus  
 Beet yellow stunt virus  
 Viruses; serNA positive-strand viruses, no DNA stage;  
 Closteroviridae; Closterovirus.  
 1 (bases 2734 to 5576)  
 Karasev,A.V., Nikolaeva,O.V., Koonin,E.V., Gumpf,D.J. and  
 Garnsey,S.M.  
 Screening of the closterovirus genome by degenerate primer-mediated  
 polymerase chain reaction:  
 J. Gen. Virol. 75 (Pt 6), 1415-1422 (1994)  
 9267425  
 9207405  
 2 (bases 1 to 10545)  
 Karasev,A.V., Nikolaeva,O.V., Mushagian,A.R., Lee,R.F. and  
 Dawson,W.O.  
 Organization of the 3'-terminal half of beet yellow stunt virus  
 genome and implications for the evolution of closteroviruses  
 Virology 221 (1), 199-207 (1996)  
 96266429  
 8661428  
 3 (bases 1 to 10545)  
 Karasev,A.V., Nikolaeva,O.V., Lee,R.F., Wisler,G.C., Duffus,C.E.  
 and Dawson,W.O.  
 Characterization of the beet yellow stunt virus coat protein gene  
 Phytopathology 88 (10), 1048-1045 (1998)  
 4 (bases 1 to 10545)  
 Karasev,A.V.  
 Direct Submission:  
 Submitted (20-MAR-1996) Alexander V. Karasev, University of  
 Florida Citrus Research and Education Center, 700 Experiment  
 Station Road, Lake Alfred, FL 33850-2299, USA  
 On Jun 25, 1996 this sequence version replaced gi:507953.  
 Location/Qualifiers  
 1..10545  
 /organism="Beet yellow stunt virus"  
 /mol\_type="genomic RNA"  
 /specific\_host="Sonchus oleraceus L."  
 /db\_xref="taxon:35290"  
 <1..2133  
 /function="replication"  
 /note="ORF1a; similar to helicases of other  
 closteroviruses"  
 /codon\_start=1  
 /product="helicase"  
 /protein\_id="AAC55658.1"  
 /db\_xref="GI:1388129"  
 /translation="SSFLCVWJWRSQYCFIGNKFKSVGDSVYKAVPTTINYKRRP  
 TLDSIFEVHNHREVORLKNFLIEGLGFINEGSGSASGEEDNFEAEATKMDQ  
 RESEGTSTPANTFVNDVAGLKIETPAORSHCVSLFKADRICKYGYDKFEPFSRKK  
 TLVSKFLNLTENAPFPMIIEADHLASNVREFFYLQELILFELFKLNRYSELE  
 VVEFRKICRAGDEARLLVYRSDSLYKGDGRRLPKMEEYEFVFTGGLLPNDKEF  
 SGNGLPHANTKFTAAANSIRNSYRNFI FENDSCR:RLYEAPPGGKTHLIASFV  
 MKHKNRILVTANKSSQVEILKINDSKREHETKTLKLFASKAERNPASDVY  
 TDVSYLWHLGTCVDLFDVCEPMHAGAVTAVNSLVPRSVFFVGDSSQ:HHIERNE  
 YDVAESDLDRLVAADRLYGVOSYRCPMDVCGMLSKHYNPVATTNVESECKSLTI  
 TEINSVDDVVASKRNTYLFLOSEKKELEKHLAKKGVATVTVHEAQDTRDVLY  
 RTFGQDAPFSSFNHINVAITRHETSLTYAVLAARNENIAAACEANSLVDFRLTP  
 HFGGVLNIDVEPVTYDNRCKASAPINSDFLDYVPGSTSLNFQ:SAEMESQ  
 PFESGANNTVYRDSAPGSGTDHQVR"  
 <2132..3514  
 /function="replication"  
 /note="ORF1b; similar to RdRp's of other closteroviruses;  
 the BYSV RdRp is presumably expressed via a +1 ribosomal  
 frameshift"  
 /codon\_start=1  
 /product="RNA replicase"  
 /protein\_id="AAC55659.2"  
 /db\_xref="GI:4559421"  
 /translation="SVVRQAIIPKRRPSQEN:LSYEARNYNFITERFSGPHEYGRA  
 MAANVLKRCFDLDKVAARSDIIATIGRGFWLRSKSPSQKSLASDLETPLNLESE

IVRFLKMWKRDAXVKLDASCLVKHPAOTIMPHRKAINALYISOCDFEFKKNRVLSCSP  
 NIIFTEMTNQKPADIASNLLGPNVFNIGEDISFYDKSDQDAFIKSPFCALYEEFGR  
 GFDFELLUSWMEGEYSERNIMDQGVFICVCKSGASNTWIGNSVTVIGILAYID  
 VSKFQALFVSGDSDILPSADEIAEDICELGFETKFLTPSPVYFCSKFLVFTGDK  
 CVFVPDYLKLLVKGASGRSLSDSELFVFEVSDRLTKKEGDERVNLTLSELVHLKYE  
 FEMGNTKALSTIHLRSNLSFSLFKVFKTGMKVYVYKAKYILKKFLGYNIEPITT  
 FGDAMFVYKE  
 3522..4343  
 /function="unknown"  
 /note="ORF2"  
 /codon\_start=1  
 /product="p30"  
 /protein\_id="AAC55660.1"  
 /db\_xref="GI:1388131"  
 /translation="MALYKYSALALIVDPDRYTSFNFNLRAIFSACDARVFGVSVV  
 GSVJASTJPTLTSVFEVQNGEFLSKRYRMDERRRITTGAEITPHSDAKLFIHQV  
 VKKO:DGNTGLLVYNGHLLGGTIVRQFDILQGCENAYADLACPSRFSEVSDVL  
 CDCPMSEIINTSDNNRSACTDAFYRRVFECDGDLDTVDSDNSRHFVEVDLRASE  
 SSRGDPDRVDVLSSGEPEKXVLNRN:MCLFLVLLLAGATIVVYVNLRSFGVS"  
 4401..4556  
 /function="unknown"  
 /note="ORF3; similar to small hydrophobic proteins of  
 other closteroviruses"  
 /codon\_start=1  
 /product="p6"  
 /protein\_id="AAC55661.1"  
 /db\_xref="GI:1388132"  
 /translation="MDCILRAFUPFGPAJVICFFIAVAAYFAFFVXNTHSQDTDVDI  
 RQEDLAG"  
 4574..6409  
 /function="unknown"  
 /note="ORF4; similar to HSP70 homologs of other  
 closteroviruses"  
 /codon\_start=1  
 /product="p66"  
 /protein\_id="AAC55662.1"  
 /db\_xref="GI:1388133"  
 /translation="MVVGVDFGTFSSVCVFNPSGRHLHFVKQNSAYIPTCLFLYSDT  
 MAMSGYDAETASLDPNVKGFFRDLKRWGCDENIEEYKSLKPKHYSTLSNFGKG  
 SRK:PTLGSYSGSVQMSGLSGLIALP:QALVKSAAIEFKCECTELIVSVPANCDMQ  
 RLFTENCUNLSGFTCVHMNEPDSAAAL:YGMANLPFDEAD:SSLKESLSPKIDPITY  
 NQTFVRSAGDMNLGRDVRAPKAK:YQMANLPFDEAD:SSLKESLSPKIDPITY  
 TVKTGDESKTVVYVSRGL:LAEIVFPVDRTIKVMKRVFELYVKMNLKADQAKASLV  
 VGSYSLPGLKSLQVDFVSECDLDPRAAAGCALYSSCLSSSPKLLVDCASH  
 NLSIPNVGESIVLPAGAPVPFVGTDRDINLASCVGSQSPV:FEGDRTKCFYKVKY  
 FSGTVPLKDLGVTGCTPRTIVTLATFVSVGVTKFTIG:SAKKVFGVGPAYDPK  
 ESVSIRITELHTDNQNRVLLALTLTKTREARQKESYSEKHLDLSGNLDTESKX  
 FNGYNEQTADVCRILLGKSVQKTLRGARVEELSYRNI:EVQTLKI"  
 6285..7889  
 /function="unknown"  
 /note="ORF5; similar to analogously positioned proteins of  
 other closteroviruses"  
 /codon\_start=1  
 /product="p61"  
 /protein\_id="AAC55663.1"  
 /db\_xref="GI:1388134"  
 /translation="WSRRPTFAGYSWGLSKFRHRYGEPEKSYLTETSMKYKPLKSESI  
 TFYDSSLITSAELPARGSTAEYEAIALIIFSDSITTKSEKLSIYRGLNQ:NNHSIY  
 AETELVDYKTIQCKT:ISAVESFMGRASAAQVEHMSLSNCSGELLNPNDFARFI  
 QLVFKDAEQAOQNTSGSVSDYLYVLCQLQYDMSKKNSAGRTFKQTYLPMIDIFDY  
 QHSLYKRPEDNLLFPIDYEIT:KEVPTMSVIDANVLSKLVLDYSY:ENKSI LALEKK  
 SWELYAPDERLLFPIDYEIT:KEVPTMSVIDANVLSKLVLDYSY:ENKSI LALEKK  
 ITSLCRDNEGIDEGALWAAFFCVGYRTARQVRVCRPDTYELDGI:FSKPIVMSGVE  
 LFPPELOKRPVDVSLRRFENGAKAGEAITVFKIGISFPPT:RLNAPSKYSYLINDYF  
 QVNSLGTPEPEKILCNIAKDVTDMCAQRISSWKAPQIQRNGEAINSAKIRTLPTN  
 TLNALSCKDLUNQAPSWMNTLTINLR"  
 7832..8515  
 /function="unknown"  
 /note="ORF6; similar to analogous BYSV coat proteins of  
 other closteroviruses"  
 /codon\_start=1  
 /product="p25"  
 /protein\_id="AAC55664.1"

```

/db_xref="GI:1398135"
/translation="MPDQPGAEIVHNANKSLIEVSESTSEKVKCKFENNEDHKTFKO
VNPILLNDELREVLCKLTELKLNKALDEDIYHHVAFLLRASFVSTSKYKYGK
YSYSDQRYTVNDKAWFPQVILASKNKPNGLRAFCAJEGMFLSVAR-GPDPAGT
RVSGKRGAPSGSEYLGADFLTSCPLMSDHDRAVALSASRNALDRSAASQ-DKKXVSL
YDFGKVVYT"
misc_feature
8551
/notes="transcription start site for coat protein
subgenomic RNA"
CDS
8617..9264
/notes="CP; CRF7; similar to coat proteins of other
cisteoviruses"
/codon_start=1
/evidence=experimental
/product="coat protein"
/protein_id="AAC55665.1"
/db_xref="GI:1398136"
/translation="VAGNDEGSDSSASQVTAKDMIFAPFENFARASATCJRGXNK
KXLFEEFSVRKTDQVTESGIPTILGMLYALATSTSKIDIEDKTLPLVSAKIDAVN
VTITVEDIKFNVSUTLLKNYKNKLRVFARTPEEYLRVFCYKHILPNIARANKHG-
PADYSYLAADFVQTSNLLKHEQAVLLEGRNAATASSGTTRESAVNLVYLGSSK"
9261..9743
/function="unknown"
/notes="ORF8"
/codon_start=1
/product="p18"
/protein_id="AAC55666.1"
/db_xref="GI:1398137"
/translation="KXVIESLQFVTPVHEKYALVLTDDAFIFEHVLGVQSSYPVL
WHGEGADMGSVISFTYKQLQVNLNLSLQYTRMLRVKRTTSSLSLEJINVRANAV
SIFRCNAPLITVTSKGPESMDAVLYNGIIGTPVCPTHLSS-VESYGNITHFAE-
"
9746..10306
/function="unknown"
/notes="ORF9; similar to the 20-kDa protein of CTV and to
the 21-kDa protein of BYV, two other closteroviruses"
/codon_start=1
/product="p22"
/protein_id="AAC55667.1"
/db_xref="GI:1398138"
/translation="MKLLISDSYVYDSINLANKLNELNKNKSVPLESIASYNELLK
VNAQSGDAIRKNEKNENLFPDLSVNTDLEMLKVRDQLIRAEINKFNCSSQDVV
KLIIRINVTNHPNSRERPEVTEQIOFSMRGVINTSLQNSYKYKFEVTINYPKCA
LLELVKNGSVKTFVYENLESEFJ"
10304..10545
/notes="similar to 3' UTR's of CTV and BYV"
3'UTR
2989 a 2196 c 2497 g 2863 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 6,95e-18 Length: 10545
Score: 266.50 Matches: 66
Percent Similarity: 48.09% Conservative: 22
Best Local Similarity: 36.07% Mismatches: 94
Query Match: 26.89% Indels: 1
DB: 14 Gaps: 1
US-09-613-486-15 (1-198) x BYUS1931 (1-10545)
QY 11 AsnLeuValIleThrAspAlaSerSerLeuAsnGluValAspIysLysLeuLeuSerAla 30
Db 8704 AATTTCGTCGCGGAGCGGACCTTGTGTAATGGTGAAGAAACAAAGAAAGTGTTCGAG 8763
QY 31 GluValGluLysMetLeuValGlnIlyGlyAlaProAsnGluGluIleGluValValPhe 50
Db 8764 GAATTTTCTGTGAGAGTGAAGACTCAGAGTGTCTAGTGAATCCGGCATCCCAACATTG 8823
QY 51 GlyLeuLeuLeuTyralaLeuAlaAlaAArgThrThrSerProLysValGlnArgAlaAsp 70
Db 8824 GGAATGACGCTGTACCCCTAGCTACCTTAAGCACTTCTCTAAATTCGATATCGAGGAT 8883
QY 71 SerAspValIlePheSerAsnSerPheGlyGluArgAsnValValValThrGluGlyAsp 90

```

```

Db 8884 AAGACACCGCTAGTAGCGGCTAAGATAGATCGGTCAACGTGACGATAAATTACGAGAC 8943
QY 91 LeuLysIysValLeuAspGlyCysAlaProLeuThrArgPheThrAsnLysLeuAArgThr 110
Db 8944 ATTAAAGAACTTCGTAAACACTCCITGACGTTCGTTAAGAACTATAAGAAATAAATCCGAGTC 9003
QY 111 PheGlyArgThrPheThrGluAlaTyralaAspPheCysIleAlaTyralaLysLysLeu 130
Db 9004 TTGCTCGCACTTTTGAGGAGGAGTGACCTACGCTTGTGAGGCAATACAAACACATCCCTC 9063
QY 131 ProGlnLeuAsnAlaAlaAlaGluLeuGlyIleProAlaGluAspSerTyralaAlaAla 150
Db 9064 CCGAATATCGAAGAGCGAAACACACGCGTATCCCGCTGATCTACTCGTATTAGCAGCT 9123
QY 151 AspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSerArgLysMetPheAla 170
Db 9124 GACTTCGTCCAACTAGCAACTTATTGAAAGAGCAGCAACAGGCA---GTACTTCTGGAA 9180
QY 171 SerMetTyralaLeuLysThrGluGlyGlyValValAlaAsnThrProValSerAsnLeuArg 190
Db 9181 GGTAGAAACGCTCTACAGCATCTTCGGCACCACCTAGGAGTCTCGGCTCAACCTATAAA 9240
QY 191 GlnLeuGly 193
Db 9241 TATCTTGGG 9249
RESULT 15
AF501867
LOCUS
DEFINITION Citrus tristeza virus coat protein gene, complete cds.
ACCESSION AF501867
VERSION AF501867.1 GI:20454248
KEYWORDS
SOURCE
ORGANISM
Citrus tristeza virus
Citrus tristeza virus
Viruses; ssRNA positive-strand viruses, no DNA stage;
Closteroviridae; Closterovirus.
REFERENCE
1 (bases 1 to 672)
AUTHORS
Roy, A., Ramachandran, P. and Brilansky, R. H.
TITLE
Grouping and comparison of Indian citrus tristeza virus isolates
based on coat protein gene sequences and restriction analysis
patterns
JOURNAL
Arch. Virol. 148 (4), 707-722 (2003)
MEDLINE
22548851
PubMed
12664295
REFERENCE
2 (bases 1 to 672)
AUTHORS
Roy, A., Ramachandran, P. and Brilansky, R. H.
TITLE
Direct Submission
Submitted (12-APR-2002) Plant Pathology, University of Florida, 700
Experiment Station Road, CREC, Lake Alfred, FL 33850, USA
FEATURES
Location/Qualifiers
1..672
/organism="Citrus tristeza virus"
/moi_type="genomic RNA"
/db_xref="taxon:12162"
/country="India: Bangalore"
1..672
/codon_start=1
/product="coat protein"
/protein_id="AA22216.1"
/db_xref="GI:20454249"
/translation="MDDETLLKLNKNKVKETKEGDDVVAAESSFGSNLHIDFLIAMDND
VRQLGQQAAALNRDLFTLKNVPLNLPDEKDPH:AMMLYRLAKVSSLSQSDDTTG
ITVREGVEVDLSKLTWTFNSKGIQNTALRVGRTNDALVAFRCQNRNLSYG
GRPLDAGIIPAGYHLCADFLTGAGLTGLECAVVIQAKEQLLKRGADEVVTVNRQLG
KFNR"
BASE COUNT 191 a 124 c 177 g 180 t
ORIGIN
Alignment Scores:
Pred. No.: 7.71e-05 Length: 672
Score: 130.50 Matches: 61
Percent Similarity: 40.00% Conservative: 27

```

```

Best Local Similarity: 27.73% Mismatches: 75
Query Match: 13.17% Indels: 57
DB: 14 Gaps: 11

JS-09-613-486-15 (1-198) x AF501867 (1-672)

QY 1 MetGluLeuMetSerAspSerAsnLeuValIleThrAspAlaSerSerLeu 20
DB 91 ATGAACCTTACATCGATCGGACTCTG-----ATAGCATGAACGATGTCGTCAGTTG 144
QY 21 AsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40
DB 145 GGAACCCCAACAGACCCCTTTGAACAGAGATTG-----TTTCCTTACTCTGAAGGCG 198
QY 41 -----AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyrAla 56
DB 199 AAGTATCCTTAACCTTGCCTCAGCAGGATAAGGACTTCCACATAGCTATGATGTTATATCGT 258
QY 57 LeuAlaAlaArgThrThrSer----- 63
DB 259 TTAGCGGTTAAGAGTTTCATCGTTGCAAGTGATGACACACCGGCGCATACATAACT 318
QY 64 -----ProLysValGlnArgAlaAsp-----SerAspValIlePhe---Ser 76
DB 319 CGGGAGGGTGT--GAAGTGGATTGTCTGACAAGCTTTGGACTGACGTCGTGTTAACTCC 378
QY 77 AsnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLysValLeuAsp 96
DB 379 AAGGTATTGGTAACCGT----- 396
QY 97 GlyCysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr 116
DB 397 -----ACTAATGCCCTTCGAGTCTGGGCTAGGACTAACGAT 432
QY 117 GluAlaTyrValAspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaIa 136
DB 433 GCCCTTTATTAGCTTCTGT---AGACAGAAATCGCAATTGAGTTATGGTGGACGTCGC 489
QY 137 AlaGluLeuGlyIleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCys 156
DB 490 CTAGATCGCAGGGGATTCGGCTGGATATCATTAATGTCAGATTTCCTG---ACCGGA 546
QY 157 ProLysLeuSerGluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLys 176
DB 547 GCTGGCTTGACTGATTAGNATCGCTGTGTACATACAGCTAAGGAACAATTTGTGAG 606
QY 177 ThrGluGlyGly-----ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArg 194
DB 607 AAGCGAGGGGCTGATGAAGTCGTA-----GTTACTAATGTCAGGAGCTTGGGAAA 657

```

Search completed: November 8, 2003, 00:59:54  
Job time : 3533 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OX protein - nucleic search, using frame\_plus\_p2n model

Run on: November 7, 2003, 23:40:55 ; Search time 267 Seconds  
(without alignments)  
2001.830 Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 991

Sequence: 1 MELMSDLSNLNVLITDASSL.....GVVNTPVSNLRQLGRREV 198

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5135512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DBV=xih  
-Q=/cgn2\_1/USPTO.spool/US09613486/runat\_07112003\_120411\_27082/app\_query.fasta\_1.39:  
-DB=N Geneseg 19Jun03 -QFMT=fastap -SUFFIX=ring -M=MMATCH=0.1 -DOFCS=0  
-LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=bicsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09613486.acgn\_1\_1\_312.grunat\_07112003\_120411\_27082 -NCPU=6 -ICPUJ3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=2.0 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseg 19Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1930.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	597	20	Grapevine leafroll
2	991	100.0	15500	20	Grapevine leafroll
3	286	29.9	688	16	Sugar beet yellows
4	285	28.8	783	16	BYV coat protein c
5	89.5	9.0	1500	24	C. albicans BAX-as
6	86	8.7	1197	21	Arabidopsis thaliana
7	86	8.7	1199	21	Arabidopsis thaliana
8	86	8.7	349980	22	Pyrococcus abyssi
9	85	8.6	4467	18	PSR800 fragment 1
10	84.5	8.5	4766	22	GLRAV-5 nucleotide
11	83.5	8.4	640681	24	Buchnera sp. genom
12	81.5	8.2	1209	21	Erysipelothrix rhu
13	81.5	8.2	1881	21	Erysipelothrix rhu
14	81.5	8.2	2040	21	Erysipelothrix rhu
15	81.5	8.2	2814	21	Erysipelothrix rhu
16	81.5	8.2	910715	20	Borrelia burgdorferi
17	81	8.2	349980	24	Bifidobacterium lo
18	80.5	8.1	10709	22	Genomic fragment #
19	80	8.1	2133	24	Streptococcus poly
20	80	8.1	2187	24	Streptococcus poly
21	80	8.1	2365589	24	Genomic sequence o
22	79	8.0	1356	23	Enterococcus faeca
23	79	8.0	1434	18	Grapevine leafroll
24	79	8.0	2572	23	Drosophila melanog
25	79	8.0	4827	23	Drosophila melanog
26	79	8.0	11597	20	Enterococcus faeca
27	79	8.0	11597	24	Enterococcus faeca
28	79	8.0	17915	21	Grapevine leafroll
29	79	8.0	61313	23	Propionibacterium
30	79	8.0	534720	19	Rhizobium species
31	79	8.0	536165	19	Rhizobium species
32	78.5	7.9	2282	22	Nucleotide sequenc
33	77.5	7.8	665	24	Bacillus clausii 9
34	77.5	7.8	4494	23	Drosophila melanog
35	77	7.8	912	22	Aspergillus nidula
36	77	7.8	929	22	Aspergillus nidula
37	77	7.8	2091	24	Candida albicans e
38	76.5	7.7	7676	19	Human galactokinase
39	76.5	7.7	8095	24	Gene #2228 used to
40	76	7.7	600	22	Human foetal liver
41	76	7.7	600	22	Probe #8204 for ge
42	76	7.7	600	22	Human brain expres
43	76	7.7	600	22	Human bone marrow
44	76	7.7	600	22	Probe #7406 for ge
45	76	7.7	600	22	Probe #1068 used

# ALIGNMENTS

RESULT 1

AAV08870

ID AAV08870 standard; cDNA; 597 BP.

XX AAV08870;

XX 29-MAR-1999 (first entry)

XX Grapevine leafroll virus type 2 coat protein ORF6 product.

XX GRVAV-2; closterovirus; grape; tobacco; transgenic plant;

XX disease resistance; virus resistance; beet yellows virus;

XX tristea virus; coat protein; ss.

XX Grapevine leafroll virus type 2.

XX WO9853055-A1.



FT	3' UTR	:5285..:15500
FT	/metaS= j	
FT	/note= "Claim 48"	
XX		
XX	MO9853055-A1.	
XX		
PD	26-NOV-1998.	
XX		
XX	20-MAY-1998;	98WO-US10313.
PF		
XX		
XX	20-MAY-1997;	97JUS-0047194.
XX	(CORR ) CORNELL RES FOUND INC.	
PA	Gonsalves D,	Ling K, Zhu H;
XX		
XX	WPI; 1999-045307/04.	
DR	P-PDB: AAW73476, AAW73477, AAW73478, AAW73479, AAW73480, AAW73481,	
DR	AAW73482, AAW73483, AAW73484.	
XX		
PT	Grapevine leafroll virus (type 2) proteins and polypeptides - and	
PT	encoding DNA, useful e.g. to impart grapevine leafroll resistance to	
PT	grape and tobacco plants and detect grapevine leafroll virus	
XX		
PS	Disclosure; Page 12-20; 151pp; English.	
XX		
CC	A total of 15,500 bp of the RNA genome of grapevine leafroll virus	
CC	type 2 (GLRAV-2) has been sequenced and cloned from GLRAV-2	
CC	isolated from infected Vitis vinifera cv. Pinot Noir. About 85%	
CC	of the total RNA genome was revealed from 2 different clones. The	
CC	sequence in the coat protein gene region was determined and	
CC	confirmed from several overlapping clones. The genomic	
CC	organisation of GLRAV-2 includes 9 open reading frames (see also	
CC	AAV08864-72) encoding a polyprotein (see AAW73476) having papain-like	
CC	protease, methyltransferase and helicase motifs, an RNA-dependent	
CC	RNA polymerase, heat shock proteins, coat proteins and other	
CC	proteins of unknown function (see AAW73476-84). These proteins can	
CC	be used to produce antibodies, useful to detect GLRAV in samples	
CC	e.g. by ELISA (claimed). The nucleic acid molecules can be used to	
CC	produce probes and primers for such detection, and to transform	
CC	host cells (especially Agrobacterium vitis, Agrobacterium	
CC	tumefaciens, grape, citrus, beet or tobacco cells) and produce	
CC	transgenic plants (claimed). In particular, they can be used to	
CC	impart GLRAV-2 resistance to Vitis scion or rootstock cultivars or	
CC	Nicotiana (claimed). Because extensive similarity exists between	
CC	nsP70-related sequence regions of GLRAV-2 and other closteroviruses,	
CC	the DNA may also be used to impart beet yellows virus resistance to	
CC	beet cultivars or tristezza virus resistance to citrus scion	
CC	cultivar/rootstock cultivars (claimed).	
XX		
SQ	Sequence 15500 BP; 3984 A; 3178 C; 3970 G; 4369 T; 0 other;	
Alignment Scores:		
Pred. No.:	8,13e-115	Length: 15500
Score:	991.03	Matches: 198
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	20	Gaps: 0
US-09-613-486-15 (1-196) X AAV038674 (1-15500)		
Oy	1 MetGluLeuMetSerAspSerAsnLeuSerAenLeuValIleThrAspAlaSerSerLeu 20	
Dd	13584 ATGGAGTTGATGTCGCACAGCACCTTACCAACCGTGGTGATTAACCGCGCCTTAGTCTA 13643	
Oy	21 AsnGlyValAspLysLysLeuLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40	
Dd	13644 AATGGTGTCGACAGAAGACTTTATTCTGCTGANGTTGAATAAATCTTGGTGCAGAACGG 13703	
Oy	41 AlaProAsnGluGlylleGluValValPheGlyLeuLeuLeuTyrtAlaLeuAlaIaArg 60	
Dd	13704 GTCTCTAACGAGGATAGAAGTGGTGTTCGCTACTCTTACGCACTTCGCCGAAGA 13763	



CC transcribed RNA isolated from purified SBV. The fragment is useful for  
CC the production of virus resistant transgenic plants by genetic  
CC engineering methods.  
CC (Updated on 23-MAR-2003 to correct PN field.)  
XX (Updated on 25-MAR-2003 to correct PF field.)  
XX  
SQ Sequence 688 BP; 190 A; 155 C; 157 G; 186 T; 0 other;

Alignment Scores:  
Pred. No.: 9.9e-27 Length: 688  
Score: 286.00 Matches: 71  
Percent Similarity: 52.91% Conservative: 29  
Best Local Similarity: 37.57% Mismatches: 87  
Query Match: 28.86% Indels: 2  
DB: 16 Gaps: 2

US-09-613-486-15 (1-198) x AA087853 (1-688)

QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26  
DB 104 GCGACTTTTGAAACGTAAGTCTCGCAGACCAACCTGTTTCACGGAGACACTGCCAT 163  
QY 27 LeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIle 46  
DB 164 AAACCTTAGGAAGAACTTCGAAGAGTGTGTAATAATTAAGAGGGTTCGGAAGATAACCTC 223  
QY 47 GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArgThrThrSerProLysVal 66  
DB 224 GGAATCGCGTTAGGACTTGTGTGTAATTCCTGCTACGATAGGCACCTCCAAACAAAGT 283  
QY 67 GlnArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal 95  
DB 284 AAGCTCAACCGACGCTACCTTCATCAAGCTTCGTTTGGTGGTGGAGGAACTGTAC 343  
QY 86 ValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105  
DB 344 CTCACCTCAGCGTGAATTAATTCCTTTCTGGGGTCTCAAAAACCTTTGGAGGAAACCT 403  
QY 106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPheCysIleAla 125  
DB 404 AACAAATTGCGGTGTTCTCGCGTACTTTTCAGAGGACTACATATCCTTGGCGCAGGAA 463  
QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145  
DB 464 TACCGAGGGAATTAACCTCCGA\*TGCCAGAGCTAACCGTCACGCTACCCGCTCAAGAT 523  
QY 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165  
DB 524 CACTACTTAGCCGCTGACTTCATATCGAGCTGACGGAACCTCACGACCTCAACAAAGT 583  
QY 166 ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrPro 185  
DB 584 CGTCTCTGTAGTCGCGCGGAAACGCCACTCACAGGAATTCCTCG---TCTGAATCACCG 640  
QY 186 ValSerAsnLeuArgGlnLeuGlyArg 194  
DB 641 GAACCAAGTTGAACCAACTAGTCTGT 667

RESULT 4

AA042975

ID AA042975 standard; cDNA; 783 BP.

XX

AC

XX

DT

DE

XX

KW

KW

KW

KW

KW

KW

KW

XX Beet yellow virus.  
OS  
XX  
FH Key Location/Qualifiers  
Ff sig\_peptide /tag= a  
Ff /note= "Omega sequence from TMV"  
Ff 89..703  
Ff /tag= b  
Ff /product= BYV\_coat\_protein  
FT  
XX W09428147-A1.  
XX 08-DEC-1994.  
XX 01-JUN-1994; 94MO-EP02786.  
XX 02-JUN-1993; 93GB-0311332.  
XX (SANO ) SANDOZ LTD.  
XX (SANO ) SANDOZ PATENT GMBH.  
XX (SANO ) SANDOZ-ERFINDUNGEN VERW GES MBH.  
XX Boisen K, Brunsted J;  
XX WP; 1995-022818/03.  
XX Recombinant DNA sequence comprising at least two coat protein genes  
XX - used to combat viruses in e.g. sugar beet  
XX  
XX Disclosure; Page 11-12; 23pp; English.  
XX The sequences given in AAT42975-77 represent cDNA sequences encoding  
XX coat proteins from beet yellow virus (BYV), beet western yellow  
XX virus, (BWV) and beet necrotic yellow vein virus (BNYVV)  
XX respectively. These sequences are used in the recombinant DNA  
XX sequence of the invention. This comprises nucleotide sequences  
XX encoding the coat proteins of at least one luteovirus and at least  
XX one closterovirus. This recombinant DNA may be used to transform  
XX sugar beet for combatting viral infection. Other plants may also  
XX be transformed e.g. fruit such as mangoes, apples, pears, bananas,  
XX and field crops such as sunflower, wheat, barley, maize, and  
XX vegetables such as potatoes, carrots, cabbage and onion.  
XX  
SQ Sequence 783 BP; 233 A; 178 C; 171 G; 211 T; 0 other;

Alignment Scores:  
Pred. No.: 1.6e-26 Length: 783  
Score: 285.00 Matches: 70  
Percent Similarity: 53.44% Conservative: 31  
Best Local Similarity: 37.04% Mismatches: 86  
Query Match: 28.76% Indels: 2  
DB: 16 Gaps: 2

US-09-613-486-15 (1-198) x AAT42975 (1-783)

QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26  
DB 104 GCGACTTTTGAAACGTAAGTCTCGCAGACCAACCTGTTTCACGGTGAAGACTCGAC 178  
QY 27 LeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIle 46  
DB 179 AAACCTAGGAAGAAATTCGAAGAGTGTGTAATAATTAAGAGGGTTCGGAAGACAAC 238  
QY 47 GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArgThrThrSerProLysVal 66  
DB 239 GGTCTCGCGTTAGAACCTTGTGTGTAATTCCTGTCGACGATAGTACTTCTATAAAGT 298  
QY 67 GlnArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal 85  
DB 299 AGTCTCCAAACCGACGCTACTTTCATCAACAAAGCTTCGTTTCGTCGTCGGAAGAA 358  
QY 86 ValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105

Db 359 CTCACTCAGGTGAAGTGGCTTTTCTGGACTCTCAGAACTTTTGGAGGAAGCCT 418  
 QY 106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPheCysIleA 125  
 Db 419 AACAAAGTTGGTTGTTCTCGCGCACTTTTCAGAGGACTACATATCCTTCGCGAAGGAA 478  
 QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyLeuProAlaGluAsp 145  
 Db 479 TACCGAGGAAGACTGCTCGGATTTGTAGAGCAACCGTCACGGTCTACTGCTGAAGAT 538  
 QY 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165  
 Db 539 CACTACTTACTGCTGATTTTCATATCAGATCAACAGAACTTACTGACCTACACAAAGT 598  
 QY 166 ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGlyValValAsnThrPro 185  
 Db 599 CGTCTGCTGTGGCGCGGAAACGGCACTTCACACAGAGTTCTCG---TCTGAATCACCA 655  
 QY 186 ValSerAsnLeuArgGlnLeuGlyArg 194  
 Db 656 GTAACCTAGTTTGAACAGCCTGGGTCTG 682

## RESULT 5

ABQ76539  
 ID ABQ76539 standard; cDNA; 1500 BP.

XX AC ABQ76539;

XX XX 21-NOV-2002 (first entry);

DE C. albicans BAX-associated cDNA fragment SEQ ID 503.

XX Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;  
 XX vasotrophic; vaccine; gene therapy; proliferative disorder; cancer;  
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;  
 KW neurodegeneration; cell death; ss.  
 XX OS Candida albicans.

XX W200264766-A2.

XX 22-AUG-2002.

XX 21-DEC-2001; 2001WO-EPI5398.

XX 22-DEC-2000; 2000EP-0870318.

PR 04-JAN-2001; 2001EP-0870002.

PR 09-JAN-2001; 2001EP-0870003.

XX (JANC ) JANSEN PHARM NV.

XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;

XX WPI; 2002-667002/71.

DR P-PSDB; ABG93273.

PT New isolated nucleic acid representing a synthetic Bax gene, useful as  
 PT medicament for treating, preventing and/or alleviating yeast or fungal  
 PT infections or proliferative disorders, or for preventing apoptosis in  
 PT certain diseases .

XX Claim 36; Figure 2; 344pp; English.

XX This invention describes a novel nucleic acid representing a synthetic  
 CC Bax gene. The Bax gene of the invention is useful for identifying  
 CC Bax-resistant yeast or fungi, identifying, or obtaining and identifying  
 CC Candida spp. sequences that are differentially expressed in a pathway  
 CC eventually leading to programmed cell death or identifying inhibitors or  
 CC inhibitor sequences of Bax-induced cell death. The products of the  
 CC invention have cytostatic, fungicide, immunosuppressive, virucide and  
 CC vasotrophic activity and can be used in vaccines or for gene therapy. The  
 CC isolated nucleic acids, polypeptides, pharmaceutical compositions,

CC antisense molecules and antibodies are useful as medicaments or in  
 CC preparing a medicament for treating, preventing and/or alleviating  
 CC diseases associated with yeast or fungi or proliferative disorders, such  
 CC as cancer, or for preventing apoptosis in certain diseases. The compounds  
 CC or polypeptides, or the genetically modified organism are useful for  
 CC preparing a medicament for modifying the endogenic flora of humans and  
 CC other mammals. The vaccine is useful for immunising against yeast or  
 CC fungal infections. Apoptosis-related diseases include autoimmune disease,  
 CC ischaemia, diseases related with viral infections or neurodegenerations.  
 CC This sequence represents a polynucleotide associated with the Bax gene  
 CC described in the disclosure of the invention.

XX SQ Sequence 1500 BP; 442 A; 291 C; 309 G; 458 T; 0 other;

## Alignment Scores:

Pred. No.: 0.387 Length: 1500  
 Score: 89.50 Matches: 53  
 Percent Similarity: 43.60% Conservative: 39  
 Best Local Similarity: 25.12% Mismatches: 76  
 Query Match: 9.03% Indels: 43  
 Ds: 24 Gaps: 12

US-09-613-486-15 (1-198) x ABQ76539 (1-1500)

QY 3 LeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSer---LeuAsn 21  
 Db 517 TTGATGGCTTCTGGAAATGGGACAGTCTTCTGCTACTGCTTCTACCACTGTTTGAAG 576  
 QY 22 GlyValAspLysLysLeuLeuSerAla---GluValGluLysMetLeuValGlnLysGly 40  
 Db 577 ACTGCTGAATCCACCCCATATCTGCTTTATCTTCCCAATTGTAGTCGAAGCCGGT 636  
 QY 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArg 60  
 Db 637 ATGCCAAAGGTGTTACACATTTGTTTCGGTTTGGTGTCTGCTGCTGCTGCTGCTGCT 696  
 QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80  
 Db 697 GCTAACATCCAAAGATTGAAAAA-----GTGCTTTTCACTGGTTCTACTGCC 744  
 QY 81 GluArgAsnValValVal-----ThrGlyAspLeuLysLysVal----- 94  
 Db 745 ACCGGTAAATTAATCATGAATTCGGTCTGAATCAAACTTGAAAAAGTTTACTTTGGAA 804  
 QY 95 LeuAspGlyCysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThr 114  
 Db 805 TTGGTGGTAAATCTCCAAACATTTTTCACCAT-----GCTGATTGCGACAAGACT 858  
 QY 115 PheThrGluAlaTyrVal-----AspPheCysIleAla----- 125  
 Db 859 ATTCAAAACCTTGATTGTTTCTATCTTACAAATTCGTGGAAGTCTGTTGCTGTTCT 918  
 QY 126 -----TyrLysHisLysLeuProGlnLeuAsnAlaAlaAla 137  
 Db 919 CGTCTTTTGAATTCGCGTGTTCAGACCAAGTGTGTGAAAAAATTCAAAGAAGCTGCT 978  
 QY 138 Glu-----LeuGlyIlePro---AlaGluAspSerTyrLeuAlaAlaAspPheLeu 153  
 Db 979 GAAAGTGTCAAGGTGTGTAACTTCGACGACACACTTTTATGGGTGCC----- 1029  
 QY 154 GlyThrCysProLysLeuSerGluLeuGlnGlnSerArgLysMetPheAlaSerMetTyr 173  
 Db 1030 -----CAAGTTTCTGACGTCCAATTTGTC-----AAAATTTTCAATACGTTGAA 1074  
 QY 174 AlaLeuLysThrGluGlyGlyValValAsnThr 184  
 Db 1075 TCTGTAATATCTCAAGGTGCTACTGTTGTTTACC 1107  
 RESULT 6  
 AAC45964  
 ID AAC45964 standard; DNA; 1197 BP.  
 XX AC AAC45964;



```

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161520.
PR 28-OCT-1999; 99US-0161592.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Alignment Scores:
Pred. No.: 0.78
Score: 86.00
Percent Similarity: 38.92%
Best Local Similarity: 26.49%
Query Match: 8.68%
DB: 21
Length: 1197
Matches: 49
Conservative: 23
Mismatches: 93
Indels: 20
Gaps: 9

```

US-09-613-486-15 (1-198) x AAC45964 (1-1197)

```

Qy 7 SerAsnLeuSerLeuValIleThrAspAlaSerLeuAsnGlyVal-----Asp 24
Db 427 TCGGAACCAACCAACATGTTCAAGAGTTACATTGCTGCTCAAGGTATCGTGGGAGAC 486
Qy 25 LysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGlu 44
Db 487 AAGAAACGATCGAGATCATAAACAATGCCTTCGTGTTGTGAGTCAGGGCCCTACCAT 546
Qy 45 GlyIle-----GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArg 60
Db 547 TTCAATTGAACTATTACGAATATTCCTCAAGGGCTGTCGAGTATCCTTTCAATTCTCGT 606

```

```

Qy 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
Db 607 TACCAAGACTTTATCTTAAGAGGCTTGAATAATTCCTGGGAGCTTTACAGTTTAGT 666
Qy 91 GluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaPro 100
Db 667 GTCCGGATGTTTGTT-----GGAGGTTTACCGCAATG-----GGTGTTTACCG 714
Qy 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrVal 120
Db 715 ATC---CATATGACTGCTAAATTCGCAACATTTTAGATTCTGCTTGGAAACCACTAAC 771
Qy 121 AspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140
Db 772 AAGACTCTGTTTATACAAATGAGAACTTCAGAACTCTTGTGCCCCAAATCGAAGCATCT 831
Qy 141 IleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
Db 832 CTTCAGGA---ACCAATTCCTTTACGCCGATGCTCTATAAT-----CCTATGATGGAG 882
Qy 161 GluLeuGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLys---ThrGluGly 179
Db 883 ATGATCCAAAACCTACGAAA-----TACGGTTTCAAGAGACCAAGAGA 927
Qy 180 GlyValValAsnThr 184
Db 928 GGTGTTGTGGACA 942

```

RESULT 7  
AAC39C12  
ID AAC39C12 standard; DNA; 1199 BP.  
XX AAC39C12;  
XX  
D: 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23034.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
EP:033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134258.  
PR 14-MAY-1999; 99US-0134259.  
PR 14-MAY-1999; 99US-0134261.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134766.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139219.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0142287.  
PR 01-JUL-1999; 99US-0142842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142190.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0143320.  
PR 12-JUL-1999; 99US-0143977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147535.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 26-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.: 0.782 Length: 1199  
 Score: 86.00 Matches: 49  
 Percent Similarity: 38.92% Conservativity: 23  
 Best Local Similarity: 26.49% Mismatches: 93  
 Query Match: 8.68% Indels: 20  
 DB: 21 Gaps: 9

US-09-613-486-15 (1-198) x AAC39012 (1-1193)

QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyVal-----Asp 24  
 DB 429 TCGGAACACCAACATGTTCAAGAGTTACATGTCGTCGTCACAGGTATCGTGGAGAC 48a  
 QY 25 LysLysLeuLeuSerIaGluValGluLysMetLeuValGluLysGlyAlaProAsnGlu 44  
 DB 489 AAGAAAGCATGGAGATGATATAACAAATGCTTCGTGGTTGTGAGTGCAGGGCTTACAGT 54a  
 QY 45 GlyIle-----GluValValPheGlyLeuLeuLeuTyAlaLeuAlaIaArg 60  
 DB 549 TCAATTTGACTATACGMAATCCCTCAAGCGTCTCGATATCCTTCATTCCTGCT 60a  
 QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80  
 DB 609 TACCAAGACTTATTCTTAAGAGCGTTGAAATTTCTGTCGGGAGCTTTACAGTTTAGGT 66a  
 QY 81 GluArgAsnValValValThrGluGlyAspPheLysLysValLeuAspGlyCysAlaPro 100  
 DB 669 GTCCGGAATGTTTGTTT-----GGAGGTATACCGCCCAATG-----GGGTGTTTACCG 716  
 QY 101 LeuThrArgPheThrAsnLysLeuValGlyArgThrPheThrGluAlaTyVal 120  
 DB 717 ATC---CATATGACTGCTAAATTCGCAACATTTTATAGATTCGCTTGGACACCATATAC 773  
 QY 121 AspPheCysIleAlaTyLysHisLysLeuProGlnLeuAsnAlaIaIaGluLeuGly 140  
 DB 774 AAGAGACTCTGTTTATACATGAGAACTTCAGAACTCTCTTGCCCAAAATCGAAGCATCT 833  
 QY 141 IleProAlaGluAspSerTyLLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160  
 DB 834 CTTCACGA---AGCAATTCCTTTACGGCGATGCTATAAT-----CCTATGATGGAG 884  
 QY 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyAlaLeuLys---ThrGluGly 179  
 DB 885 ATGATCAAAACCCCTAGCAAA-----TACGGGTTCAAGAGACGAAGACA 929  
 QY 180 GlyValValAsnThr 184  
 DB 930 GGATGTTGTGGAACA 944

## RESULT 8

AAP86431

ID AAP86431 standard; DNA; 349980 BP.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

KW Hyperthermophilic archaeon; hyperthermophilic protein; ds.

OS Pyrococcus abyssi.

PH Key Location/Qualifiers

FT misc\_feature 300001..349980

PT /\*tag= a

PT /notes= "This sequence overlaps with the 5' end of AAH41223"

XX FR2792651-A1.

XX 27-OCT-2000.

XX 21-APR-1999; 99FR-0005034.

XX 21-APR-1999; 99FR-0005034.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX (IPRE-) IPREMER INST FR RECH EXPL MER.

XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;

XX Querellou J, Weissenbach J, Saurin W, Heilig R;

XX WPI; 2001-126236/14.

XX New nucleotide sequences isolated from Pyrococcus abyssi encode

XX proteins useful in industry -

XX Claim 1; Page 183-279; 1657pp; French.

XX The present invention relates to the genomic sequence of Pyrococcus

XX abyssi and P. abyssi proteins (see AAB96053-AAB96842). P. abyssi is a

XX hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal

XX vents. The present sequence is a fragment of the genomic sequence of P.

XX abyssi. The 3' end of this sequence overlaps with the 5' end of AAH41223.

XX The proteins of the present invention have various potential industrial

XX uses, since the proteins are stable at very high temperatures, some up to

XX 110 degrees centigrade.

XX Note: This patent is in the same patent family as WO200065062, which

XX contains additional sequences as shown in AAB99132-AAB99143,

XX AAH75903-AAH75920 and AAG66436.

XX SQ Sequence 349980 BP; 95167 A; 78444 C; 78870 G; 97499 T; C other;

Alignment Scores:

Pred. No.: 2.99e+03 Length: 349980

Score: 86.00 Matches: 40

Percent Similarity: 35.71% Conservativity: 15

Best Local Similarity: 25.97% Mismatches: 41

Query Match: 8.68% Indels: 58

DB: 22 Gaps: 8

US-09-613-486-15 (1-198) x AAP86431 (1-349980)

QY 26 LysLeuLeuSerAlaGluValGluLysValLeuValGlnLysGlyAlaProAsnGluGly 45

DB 138056 AAATTGGAGACTTCATAGTTGAAAAAATGGCTGAAAGAAA---GTTCTGGGATAAGC 138112

QY 46 IleGluValValPheGlyLeuLeuLeuTyAlaLeuAlaIaArgThrSerProLys 65

DB 138113 ATTACATAATT----- 138124

QY 66 ValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGlyGluArgAsnVal--- 84

DB 138125 -----AAAGATGGGACGTTGTTCTATGCAAAAGGCTTTGGATACAGAAATGTTGAG 138175

QY 84 ----- 84

DB 138176 GCAAGGCTACCATCTACCCCGAGACAAATATACGAATTGGTCAATACCAAGAGCTTT 138235

QY 85 -----ValValThrGluGlyAspLeuLysValLeuAspGly 97

DB 85 -----ValValThrGluGlyAspLeuLysValLeuAspGly 97

```

Ddb 138236 ACAGCCTTGGCAATGTAATAGTGGAGAGGCGGNTTAAGT-----TTGGATGAC 138289
Qy 98 CysAlaProLeuThrArgPheThrAsn---LysLeuArgThrPheGlyArgThrPheThr 116
Ddb 138290 -----CCGTTGAGAAATTCGTTAATATAGCTTAGACCATTCGGAGAACCAAGTGACC 138343
Qy 117 G'uaLaTyRValAspPheCys::LeaLaTyRlySHisLysLeuProGlnLeuAsnAlaLa 136
Ddb 138344 -----GTTCAACACCTGTAACGCAATTCCTCGGGGATTCACATCTTAGGATA-GCC 138394
Qy 137 AlaGluLeu-----GlyLeuProAlaGluAspSerTyRLeu 148
Ddb 138395 GAGGCTTCATAGACGGAATGCTCGTGGGATAACTGGTTG 138436

RESULT 9
AAT68648
ID AAT68648 standard; DNA; 4467 BP.
XX
AC AAT68648;
Dc 19-AUG-1997 (first entry)
XX
DE psrQ800 fragment including coding sequence for AbiE.
XX
KW AbiE; phage abortive infection protein; phage resistance; psrQ800;
KW lactic acid bacterium; ss.
XX
OS Lactococcus lactis subsp. lactis isolate W1.
FF
FH Key Location/Qualifiers
FT -35_signal 1203..1207
FT /tag= a
FT -10_signal 1225..1230
FT /tag= b
FT RBS 1265..1271
FT /tag= c
FT CDS 1279..3078
FT /tag= d
XX
W0972C917-A2.
XX
XX 12-JUN-1997.
XX
XX 20-NOV-1996; 96MO-IB01385.
XX
XX 01-DEC-1995; 95US-0565907.
XX
XX (UNIL) QUEST INT BV.
XX
XX Holter BJ, Kondo JK, Moineau S, Vandenbergh PA;
XX Vedamuthu ER;
XX
XX WPI: 1997-3:9765/29.
XX P-PSDB; AAU17788.
XX
XX Isolated DNA encoding the AbiE protein of Lactococcus - for
XX protecting strains used in production of fermented dairy products
XX
XX Claim 4; Page 29-32; 49pp; Eng:sh.
XX
XX The 4.5 kb EcoRI fragment (AAT68648) of plasmid psrQ800 includes
XX a coding sequence for AbiE (AAU17788), a protein which, in
XX Lactococcus lactis (L1), increases resistance to phages by
XX aborting infection. psrQ800 was isolated through examination of
XX the phage resistance mechanism of L1 subsp. lactis W1. The total
XX plasmid DNA of W1 was co-electroporated with shuttle vector pSA3
XX into phage-sensitive, plasmid-free L. lactis LM0230. Phage-
XX resistant transformants were isolated and found to contain psrQ800.
XX The isolated DNA can be used to impart phage resistance to
XX bacteria, esp. L1., for use in the prodn. of fermented dairy
XX products.
XX
XX Sequence 4467 BP; 1583 A; 668 C; 643 G; 1573 T; 0 other;

```

```

Alignment Scores:
Pred. No.: 7.09 Length: 4467
Score: 85.00 Matches: 36
Percent Similarity: 41.89% Conservative: 26
Best Local Similarity: 24.32% Mismatches: 46
Query Match: 8.58% Indels: 40
DB: 16 Gaps: 6

US-09-613-486-15 (1-198) x AAT68648 (1-4467)
Qy 2 GluLeuMetSerAspSerAsnLeu-----SerAsnLeuValIleThrAsp--- 16
Ddb 2059 GAATTTTAAATGAATTTAATCTAATCTCTCGAGAAATAACTTAATTAATGATAT 2118
Qy 16 ----- 16
Ddb 2119 AAAACGAAAGTTGACAAATTCGCCGTTTGTGTATAAATCGAGTAAATCGGATATTTTCT 2178
Qy 17 -----AlaSerSerLeuSerGlyValAspLys-----LysLeuSerAla 30
Ddb 2179 TTTTGTGAAATATTACTTCACTAATTCACACGACAGCTGGATTAAAGAAATAAGCAAT 2238
Qy 31 GluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGly---IleGluValVal 49
Ddb 2239 TTTATAGATTATTGCTGAATGAAGACATTTAGGCAATAGGAGCTATAAAATGTATT 2298
Qy 50 PheGlyLeuLeuLeuTyRAlaLeuAlaAlaArgThrSerProLysValGlnArgAla 69
Ddb 2299 TTCCCGATTATA-----ACAAATACATTCGAAACAAAAAAGTA 2337
Qy 70 AspSerAspValIlePheSerAsnSerPheGlyGluArgAsnValValValThrGluGly 89
Ddb 2338 GATCTATAAAATATA--GACATACTCTTTTCGAAAGAAACATGGTTACCAATTTTAAT 2394
Qy 50 AspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThrAsnLysLeuArg 109
Ddb 2395 GTTTTCGAAATAATATTAGATTATCATTAAGAAATCAAGATTAACTAATAAGTTTTTG 2454
Qy 110 ThrPheGlyArgThrPheThrGlu 117
Ddb 2455 ACTTCTTTGAAATAATTAATGAA 2478

RESULT 10
AAA9:259
ID AAA91259 standard; DNA; 4766 BP.
XX
AC AAA91259;
XX
XX 08-MAY-2001 (first entry)
XX
XX GLRAV-5 nuc.leotide sequence.
XX
XX GLRAV-5; grapevine leafroll virus; GLRAV infection; GLRAV coat protein;
XX GLRAV HSP70 homologue protein; viral gene mapping;
XX plant disease resistance; ds.
XX
XX Grapevine leafroll virus.
XX
XX Key Location/Qualifiers
XX CDS 1..159;
XX /tag= a
XX /product= "HSP70 homologue protein"
XX /partial
XX /note= "No start codon given; Specifically claimed
XX region"
XX
XX CDS 1574..2989
XX /tag= b
XX /product= "ORF 2 protein"
XX
XX FT CDS 3285..4094
XX /tag= c
XX /product= "coat protein"

```





```
QY      85 ValValThrGluGlyAspLeuLysLysValLeuAsp----- 96
Db      435855 TTAATGAAATGGAAATATTAATAAATAATATGCAAGATAAACTTAATGCCGCTTAATG 435796
QY      97 -----GlyCysAlaProLeuThrArg 103
Db      435795 GGTGTTAAGTCTACCGGAAATGGTCGTCTGTAATCTTATTCTGCTCTACCTATGCTCTGT 435736
QY      104 PheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPheCys 123
Db      435735 ATGACTAAT-----ACCTATATG 435718
QY      124 LeAlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaAgluLeuGlyIleProAla 143
Db      435717 TTATCTGGAAGCTCTAAATAGATGATATAATTAAGCGCTCGACTATGGAATA----- 435664
QY      144 GluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGln 163
Db      435663 -----TATGCTGTAAATTTTCTGGA-----GGACAGGTAGATATTACT 435625
QY      164 GlnSerArgLysMetPheAlaSerKetTyrAlaLeuLysThrGluGlyGlyValValAsn 183
Db      435624 TCGGAAAAATTTGTATTTCTACTTCGGAGCTTATTATTAATAAATAATGGAAAAATTTCT 435565
QY      184 ThrProValSerAsnLeuArgGlnLeuGlyArg---ArgGluValMet 198
Db      435564 ACACCAATTAATAAACACTACTCTCATAGATCAGGATTAGAGTTATG 435517

RESULT 12
AAAT72313
ID AAA72313 standard; DNA; 1209 BP.
XX
XX AAA72313;
XX
DT 11-DEC-2000 (first entry)
DE
DE Erysipelothrix rhusiopathiae P2A fragment DNA.
XX
XX Protective polypeptide antigen; PPA; Fujisawa strain; swine erysipelas;
XX recombinant production; immunogenic; subunit vaccine;
XX transmucous administration; ds.
XX
XX Erysipelothrix rhusiopathiae.
OS
OS Key Location/Qualifiers
FT CDS 1..1209
FT /tag= a
FT /product= "Erysipelothrix rhusiopathiae 46.5 kD PPA
FT /note= "No initiation or termination codon given in the
FT specification"
CP3072345-B1.
XX
XX 31-JUL-2000.
XX
XX 31-MAR-1999; 99JP-0094004.
XX
XX 31-MAR-1999; 99JP-0094004.
XX
XX (NORQ ) NORINSUSANSHO KACHIKU EISEI SHIKENJUCHO.
XX (HGET ) HIGETA SHOYU KK.
XX (FUJI-) FUJITA SHOUJI.
XX
XX WPI; 2000-551432/51.
XX P-PSDB; AAB14800.
XX
XX Novel recombinant protective polypeptide antigen useful as a vaccine
XX for protecting livestock against Erysipelothrix rhusiopathiae
XX infection -
XX
XX Claim 3; Fig 4A-4C; 23pp; Japanese.
XX
```

```
CC This sequence represents DNA encoding a 46.5 kD immunogenic fragment of
CC the protective polypeptide antigen (PPA) of the bacterium Erysipelothrix
CC rhusiopathiae. This bacterium infects livestock, particularly pigs
CC (swine erysipelas), and is also able to cause disease in humans. The
CC recombinant produced PPA fragment may be used as a subunit vaccine for
CC stimulating an immune response in animals against Erysipelothrix
CC rhusiopathiae. As the vaccine can be produced recombinantly, it can be
CC produced on a large scale. The vaccine is suitable for transmucous
CC administration.
XX
```

SQ Sequence 1209 BP; 456 A; 177 C; 237 G; 339 T; 0 other;

#### Alignment Scores:

Pred. No.:	2.97	Lenath:	1209
Score:	81.50	Matches:	33
Percent Similarity:	40.60%	Conservative:	21
Best Local Similarity:	24.81%	Mismatches:	50
Query Match:	8.22%	Indels:	29
DB:	21	Gaps:	5

US-09-613-486-15 (1-198) x AAA72313 (1-1209)

Qy	60	ArgThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPhe	79
Db	358	AGATTAACGATTCCTGACCTTGATGAAGCATACGTGATTACTCTGTAATTACGAGGTG	417
Qy	80	GlyGluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAla	99
Db	418	AAACACCGTATTTAGTAAATATGAAGTAAGTTAAA-----GGTAGAGCT	465
Qy	100	ProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr	116
Db	466	CCCTTAGAACCACTTATAGTTCCTCTACAGATAGAAATTCGTAGTATGAATGAATTCGT	525
Qy	117	-----GluAlaTyrValAspPheCysIleAla-----	125
Db	526	GCAGAAGTAAATATTATTACCTGAAGCGCATGAGGATTTCTTAGTTTCAGATTCAAGCGAG	585
Qy	126	TyrLysHisLysLeuProGlnLeuAsnAlaAlaAgluLeuGlyIleProAlaGluAsp	145
Db	556	TATAATGACAACTAAATAATATCAACTTTGCTTTGGGTCTAGGGTCTAGCGAGTTTAT	645
Qy	146	SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnSer	165
Db	646	GACTAT-----AACCGGCTCGAAAAATATGATGGAATA	678
Qy	166	-----ArgLysMetPheAlaSerMetTyrAlaLeuLys	176
Db	679	GAACCTCATCCACTGATCTTGAACCTTATGCTATGCGG	717

#### RESULT 13

```
AAAT72319
ID AAA72318 standard; DNA; 1981 BP.
XX
XX AAA72318;
XX
DT 11-DEC-2000 (first entry)
DE
DE Erysipelothrix rhusiopathiae full-length PPA coding sequence.
XX
XX Protective polypeptide antigen; PPA; Fujisawa strain; swine erysipelas;
XX recombinant production; immunogenic; subunit vaccine;
XX transmucous administration; ds.
XX
XX Erysipelothrix rhusiopathiae.
OS
XX Key Location/Qualifiers
FT CDS 1..1881
FT /tag= a
FT /product= "Erysipelothrix rhusiopathiae full-length PPA"
FT mat_peptide 88..1293
FT /tag= b
FT /product= "Erysipelothrix rhusiopathiae 46.5 kD PPA
FT
```

fragment (AAB14800) "

```

FT XX JP3072345-B..
PN XX
XX 31-JUL-2000.
PD XX
XX 31-MAR-1999; 99CF-0694004.
XX 31-MAR-1999; 99CF-0694004.
PR XX
XX (NORQ) NORINSUSANSHO KACHIKU E-SEI SHIKENCOCCHO.
PA (HGET) HIGETA SHOUYU KK.
PA (FUJ-) FUJITA GAKUEN.
XX
XX WPI: 2000-551432/51.
DR P-PSDB; AAB14800.
DR
XX
XX Novel recombinant protective polypeptide antigen useful as a vaccine
XX for protecting livestock against Erysipelothrix rhusiopathiae
XX infection -
XX
XX Example 1; Fig 2; 23pp; Japanese.
XX
XX The invention relates to a 46.5 kD immunogenic fragment (AAB14800) of
XX the protective polypeptide antigen (PPA) of the bacterium Erysipelothrix
XX rhusiopathiae, and to DNA encoding it (AAV72313). This bacterium infects
XX livestock, particularly pigs (swine erysipelas), and is also able to
XX cause disease in humans. The recombinantly produced PPA fragment may be
XX used as a subunit vaccine for stimulating an immune response in animals
XX against Erysipelothrix rhusiopathiae. As the vaccine can be produced
XX recombinantly, it can be produced on a large scale. The vaccine is
XX suitable for transcutaneous administration. The present sequence represents
XX the coding sequence of DNA encoding the full-length PPA.
XX
XX Sequence 1881 BP; 719 A; 291 C; 384 G; 487 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 5.65 Length: 1881
XX Score: 81.50 Matches: 33
XX Percent Similarity: 40.60% Conservative: 21
XX Best Local Similarity: 24.81% Mismatches: 50
XX Query Match: 8.22% Indels: 29
XX DB: 21 Gaps: 5
XX
XX US-09-613-486-15 (1-198) x AAV72313 (1-1881)
XX
XX QY 60 ArgThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPhe 79
XX
XX Db 445 AGATTACCGATTCTCGAAGTGTGATGAGCATACGTTGATTACTCGTGAATTACGAGGTG 504
XX
XX QY 80 GlyGluArgAsnValValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAla 99
XX
XX Db 505 AAACACCGTATTTAGTAAATATAGAGTAAGTTAA-----GGTAGAGT 552
XX
XX QY 100 ProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr-----116
XX
XX Db 553 CCCTTAGACGATTATTAGTCTCTTAAGAGATAGAAATCGTAGTGAATGAATTGCT 612
XX
XX QY 117 -----GluAlaIleValAspPheCysIleala-----125
XX
XX Db 613 GCAGAGTAATATTATTACCTGAAGCGCATGAGGATTTCTTAGTTTCAGATTCAACGAG 672
XX
XX QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
XX
XX Db 673 TATAATCACAACTAAATAATATCAACTTTGCTTTGGGTCTAGGGTTCAGCGGTTATT 732
XX
XX QY 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
XX
XX Db 733 GACTAT-----AACCGGCTCGAAATATGATGGGAAAA 765
XX
XX QY 166 -----ArgLysMetPheAlaSerMetTyrAlaLeuLys 176
XX
XX Db 766 GAATTCATCCACTGTATTGAACTTTATGCTATCGG 804

```

```

RESULT 14
AAAS0205
ID AAAS0205 standard; DNA; 2040 BP.
XX
XX AC AAAS0205;
XX
XX C7-NCV-2000 (first entry)
XX
XX Erysipelothrix rhusiopathiae erysipelas protective antigen gene.
XX
XX Erysipelas protective antigen; Epa; SpaA.1; DNA vaccine; infection;
XX immuno-protective epitope; ds.
XX
XX Erysipelothrix rhusiopathiae.
XX
XX Key Location/Qualifiers
XX -35_signal 41..46
XX /*tag= a
XX -10_signal 65..70
XX /*tag= b
XX RBS 86..92
XX /*tag= c
XX CDS 100..1980
XX /*tag= d
XX sig_peptide 100..186
XX /*tag= e
XX mat_peptide 187..1977
XX /*tag= f
XX
XX W020C047744-A1.
XX
XX 17-AUG-2000.
XX
XX 10-FEB-2000; 2000MO-US03789.
XX
XX 10-FEB-1999; 99US-0119389.
XX (UYRQ) UNIV ROCKEFELLER.
XX
XX Fischetti VA, Shimoji Y;
XX
XX WPI: 2000-524541/47.
XX
XX P-PSDB; AAY95782.
XX
XX Vaccines for protecting turkeys and pigs against Erysipelothrix
XX rhusiopathiae infections comprising a polypeptide sequence from the
XX N-terminal region of an erysipelas protective antigen -
XX
XX Example; Fig 2; 61pp; English.
XX
XX The present sequence is that of the Epa gene of strain Fujisawa of
XX Erysipelothrix rhusiopathiae, the causative agent of erysipelas in
XX animals and erysipeloid in humans. The gene encodes a protective
XX antigen (see AAY95782), termed erysipelas protective antigen (Epa or
XX SpaA.1). It was isolated from a genomic DNA library of strain
XX Fujisawa by immunoscreening using E. rhusiopathiae convalescent
XX pig serum. The N-terminal portion of the Epa protein, especially a
XX polypeptide comprising residues 12-195 of the protein, was
XX identified as a vaccine antigen. The N-terminal polypeptide was
XX shown to protect mice and pigs from a lethal challenge with E.
XX rhusiopathiae. Vaccines containing immunogenic polypeptides
XX of E. rhusiopathiae, where the immunogenic polypeptide comprises
XX an immuno-protective epitope from the N-terminal region of Epa,
XX are claimed. A claimed method for protecting an animal,
XX especially a turkey or pig, from infection by E. rhusiopathiae
XX involves administering the vaccine, or an expression vector
XX comprising a nucleic acid encoding the N-terminal region or
XX full-length Epa protein.
XX
XX Sequence 2040 BP; 776 A; 293 C; 418 G; 553 T; 0 other;
XX
XX Alignment Scores:

```

```
Pred. No.: 6.36 Length: 2040
Score: 81.50 Matches: 33
Percent Similarity: 40.60% Conservative: 21
Best Local Similarity: 24.81% Mismatches: 50
Query Match: 8.22% Indels: 29
DB: 21 Gaps: 5

US-09-613-486-15 (1-198) x AAA50205 (1-2040)
Qy 60 ArgThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPhe 79
Db 544 AGATTACGATTCTCGAATCGATGAAGCATACGTTGATTCTCTGTAATTACGAGGTG 603
Qy 80 GlyGluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAla 99
Db 604 AACACCGGTATTTAGTAAATAATGAAGGTAAGGTTAA-----GGTAGAGCT 651
Qy 100 ProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr----- 116
Db 652 CCCTTAGAGCATTTATAGTTCCTCTAAGAGATAGAAATTCGTAGTATGAATGAATGCT 711
Qy 117 -----GluAlaTyrValAspPheCysIleAla----- 125
Db 712 GCAGAGTAATATATTACTCGAAGCGCATGAGGATTTCTTAGTTTCAGATTCAACGGAG 771
Qy 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
Db 772 TATAATGACAACTAAATAATATCACTTTGGTCTAGGGTCAGCGAGTTATT 831
Qy 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
Db 832 GACTAT-----AACCGGCTCGAAATAATGATGGAAAAA 864
Qy 166 -----ArgLysMetPheAlaSerMetTyrAlaLeuLys 176
Db 865 GAACCTTCATCCACTGTACTCTGAACTTTTGCTATGCGG 903

RESULT 15
AAA72317
ID AAA72317 standard; DNA; 2814 bp.
XX
XX AAA72317;
XX
XX 11-DEC-2000 (first entry)
XX
XX Erysipelothrix rhusiopathiae full-length PPA DNA.
XX
XX Protective polypeptide antigen; PPA; Fujisawa strain; swine erysipelas;
KW recombinant production; immunogenic; subunit vaccine;
KW transnucous administration; ds.
XX
XX Erysipelothrix rhusiopathiae.
XX Key Location/Qualifiers
XX CDS 400..2280
XX /*tag= a
XX /product= "Erysipelothrix rhusiopathiae full-length PPA"
XX /product= 487..1692
XX /*tag= b
XX /product= "Erysipelothrix rhusiopathiae 46.5 kD PPA
XX fragment (AAB14800)."
XX
XX JP3072345-Bi.
XX
XX 31-JUL-2000.
XX
XX 31-MAR-1999; 99JP-0034004.
XX
XX 31-MAR-1999; 99CP-0034004.
XX
XX (NORQ ) NORINSUSANSHO KACHIKU EISEI SHIKENJOCHO.
XX (HGET ) HIGETA SHOYU KK.
XX (FUJI-) FUJITA GAKUEN.
```

```
XX WPI; 2000-SS1432/51.
DR P-PSD3; AAB14800.
XX
XX Novel recombinant protective polypeptide antigen useful as a vaccine
PT for protecting livestock against Erysipelothrix rhusiopathiae
PT infection
XX
XX Example 1; Fig 1A-1B; 23pp; Japanese.
PS
XX The invention relates to a 46.5 kD immunogenic fragment (AAB-4800) of
CC the protective polypeptide antigen (PPA) of the bacterium Erysipelothrix
CC rhusiopathiae, and to DNA encoding it (AAA72313). This bacterium infects
CC livestock, particularly pigs (swine erysipelas), and is also able to
CC cause disease in humans. The recombinantly produced PPA fragment may be
CC used as a subunit vaccine for stimulating an immune response in animals
CC against Erysipelothrix rhusiopathiae. As the vaccine can be produced
CC recombinantly, it can be produced on a large scale. The vaccine is
CC suitable for transnucous administration. The present sequence represents
CC DNA encoding the full-length PPA.
XX
XX Sequence 2814 BP; 1048 A; 409 C; 558 G; 799 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 10.1 Length: 2814
Score: 81.50 Matches: 33
Percent Similarity: 40.60% Conservative: 21
Best Local Similarity: 24.81% Mismatches: 50
Query Match: 8.22% Indels: 29
DB: 21 Gaps: 5

US-09-613-486-15 (1-198) x AAA72317 (1-2814)
Qy 60 ArgThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPhe 79
Db 844 AGATTACGATTCTCGAATCGATGAAGCATACGTTGATTCTCTGTAATTACGAGGTG 903
Qy 80 GlyGluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAla 99
Db 904 AACACCGGTATTTAGTAAATAATGAAGGTAAGGTTAA-----GGTAGAGCT 951
Qy 100 ProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr----- 116
Db 952 CCCTTAGAGCATTTATAGTTCCTCTAAGAGATAGAAATTCGTAGTATGAATGAATGCT 1011
Qy 117 -----GluAlaTyrValAspPheCysIleAla----- 125
Db 1012 GCAGAGTAATATATTACTCGAAGCGCATGAGGATTTCTTAGTTTCAGATTCAACGGAG 1071
Qy 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
Db 1072 TATAATGACAACTAAATAATATCACTTTGGTCTAGGGTCAGCGAGTTATT 1131
Qy 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
Db 1132 GACTAT-----AACCGGCTCGAAATAATGATGGAAAAA 1164
Qy 166 -----ArgLysMetPheAlaSerMetTyrAlaLeuLys 176
Db 1165 GAACCTTCATCCACTGTACTCTGAACTTTTGCTATGCGG 1203
```

Search completed: November 8, 2003, 01:41:31  
Job time : 383 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 7, 2003, 23:44:24 ; Search time 63 Seconds  
(without alignments)  
1387.204 Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 991

Sequence: 1 XELMSDLSNLNIVITDASSL.....GGVNTVPVSNLRQLGRVYM 198

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model; -DBV=xih

-Q=/cgn2\_1/USP10.spool/US9613486/runat\_c7112003\_120411\_27115/app\_query.fasta\_1.391

-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -XINMATCH=C.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BL0SUM62 -TRANS=human40.cdi

-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORV=ext -HEAPSIZ=500 -MINLEN=C -MAXLEN=200000000

-USER=US09613486 @CNG 1 1 56 @runat 07112003 120411 27115 -NCRU=6 -CFPU=3

-NO\_MMAP -LARGEQUERY -NEG\_SCORES=C -WAIT -DSPBLOCK=100 -LONGLOG

-DBV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2\_6/prodata/2/ina/5A COMB seq:\*

2: /cgn2\_6/prodata/2/ina/5B COMB seq:\*

3: /cgn2\_6/prodata/2/ina/6A COMB seq:\*

4: /cgn2\_6/prodata/2/ina/6B COMB seq:\*

5: /cgn2\_6/prodata/2/ina/6C COMB seq:\*

6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	991	100.0	597	3	US-09-080-983-14
2	991	100.0	15500	3	US-09-080-983-1
3	86	8.7	2034	4	US-09-252-991A-1531
4	85	8.6	4467	1	US-08-565-907A-1
5	85	8.6	4467	2	US-08-910-551B-1
6	85	8.6	4467	2	US-08-909-425A-1
7	81	8.2	1743	4	US-09-252-991A-13408
8	81	8.2	2139	4	US-09-252-991A-13503
9	81	8.2	2433	4	US-09-252-991A-13950
10	79	8.0	1434	2	US-08-770-544-15
11	79	8.0	1434	4	US-09-579-259-15
12	79	8.0	536165	4	US-09-214-808-1

13	77.5	7.8	1233	4	US-09-252-991A-3735	Sequence 3735, Ap
14	77.5	7.8	1491	4	US-09-252-991A-3635	Sequence 3636, Ap
15	77.5	7.8	4884	4	US-09-252-991A-3824	Sequence 3824, Ap
16	77	7.8	1776	4	US-09-252-991A-15987	Sequence 15987, A
17	77	7.8	1794	4	US-09-252-991A-16086	Sequence 16086, A
18	77	7.8	1878	4	US-09-252-991A-16516	Sequence 16516, A
19	76.5	7.7	7676	1	US-08-451-777A-7	Sequence 7, Appl
20	76.5	7.7	7676	2	US-08-451-778A-7	Sequence 7, Appl
21	76.5	7.7	7676	2	US-08-998-208-7	Sequence 7, Appl
22	76.5	7.7	7676	5	PCT-US95-06743-7	Sequence 7, Appl
23	76	7.7	672	3	US-09-080-983-12	Sequence 12, Appl
24	75	7.6	1664976	4	US-08-916-421B-1	Sequence 1, Appl
25	74.5	7.5	1764	4	US-09-107-532A-2751	Sequence 2751, Ap
26	74	7.5	1173	3	US-09-285-601-1	Sequence 1, Appl
27	74	7.5	3515	1	US-08-596-985-1	Sequence 1, Appl
28	74	7.5	6709	3	US-09-285-601-3	Sequence 3, Appl
29	73.5	7.4	2504	1	US-08-484-105-15	Sequence 15, Appl
30	73.5	7.4	2504	1	US-08-484-106-15	Sequence 15, Appl
31	73.5	7.4	2680	3	US-08-949-153-50	Sequence 50, Appl
32	73.5	7.4	2880	4	US-09-819-964-50	Sequence 50, Appl
33	73.5	7.4	3260	4	US-09-221-017B-980	Sequence 980, App
34	73.5	7.4	4403765	3	US-09-103-840A-2	Sequence 2, Appl
35	73.5	7.4	4411529	3	US-09-103-840A-1	Sequence 1, Appl
36	73	7.4	49136	3	US-09-422-869-1	Sequence 1, Appl
37	73	7.4	59665	4	US-09-813-817-3	Sequence 3, Appl
38	73	7.4	59665	4	US-09-978-197-3	Sequence 3, Appl
39	72.5	7.3	3355	4	US-09-221-017B-412	Sequence 412, App
40	72.5	7.3	7172	4	US-08-576-165-3	Sequence 120, App
41	72	7.3	2746	2	US-08-453-702B-212	Sequence 3, App-1
42	72	7.3	7425	4	US-09-453-702B-212	Sequence 212, App
43	71.5	7.2	334	3	US-09-060-756-261	Sequence 261, App
44	71.5	7.2	334	4	US-09-670-314-261	Sequence 261, App
45	71.5	7.2	1386	4	US-09-252-991A-15531	Sequence 15531, A

#### ALIGNMENTS

##### RESULT 1

US-09-080-983-14

Sequence 14, Application US/09080983

Patent No. 6197948

GENERAL INFORMATION:

APPLICANT: Zhu, Hai-Ying

APPLICANT: Ling, Kai-Shu

APPLICANT: Gopalves, Dennis

TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS

TITLE OF INVENTION: AND THEIR USES

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP

STREET: Clinton Square, P.O. Box 1051

CITY: Rochester

STATE: New York

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER:

APPLICATION NUMBER: US/09/080,983

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA: US 60/047,194

APPLICATION NUMBER: 30,727

FILING DATE: 20-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.

REGISTRATION NUMBER: 19603/1631

REFERENCE/DOCKET NUMBER: 19603/1631

TELEPHONE: (716) 263-1304

TELEFAX: (716) 263-1603  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 597 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA

US-09-080-983-14

Alignment Scores:  
 Pred. No.: 1,016-127 Length: 597  
 Score: 991.00 Matches: 198  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-613-486-15 (1-198) x US-09-080-983-14 (1-597)

```

QY 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20
DB 1 ATGGAGTGTGATGCCGACGACCACTTACCACTCTGTGATTAACCGCGCTCTAGTCTA 60
QY 21 AsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40
DB 61 AATGGTGTCCACAGAAGCTTTTATCTGCTGAAGTTGAAAAAATGTTGGTGCAGAAAGG 120
QY 41 AATGCTAAGCGGGTATAGAGTGTGTTGGGTCTACTCCCTTACGCACCTCGCGCAAGA 180
DB 121 GCTCCTAAGCGGGTATAGAGTGTGTTGGGTCTACTCCCTTACGCACCTCGCGCAAGA 180
QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 90
DB 181 ACCACGTCCTTAAAGGTTTATCTGCTGAAGTTGAAAAAATGTTGGTGCAGAAAGG 13703
QY 81 GluArgAsnValValThrGluGlyAspLysLysValLeuAspGlyCysAlaPro 100
DB 241 GAGAGGAATGTGTAGTAAACAGAGGGTGACCTTAAAGAGGTACTCGAGGGGTGCGGCT 300
QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaVal 120
DB 301 CTCACCTAGGTTCACTAATAAATCTTAAAGGTTTGGTGTGCTACTCTTCACTGAGGCTTACGTT 360
QY 121 AspPheCysIleAlaTyrrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140
DB 361 GACTTTGTATCGGTATAGCAAAATTAACCCCAACTCAACCGCGCGCGGAATTGGGG 420
QY 141 IleProAlaGluAspSerTyrrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
DB 421 ATCCAGCTGAAGATTCGTACTTACCTGAGGATTTTCTGGGTACTTGGCCGAAGCTCTCT 480
QY 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrrAlaLeuLysThrGluGly 180
DB 481 GAATTACAGCAAGTAGGAGATGTTCCGGAGTATGTACGCTCTTAAAACTGAAGGTGGA 540
QY 181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgArgGluValMet 198
DB 541 GTGGTAATACACCAAGTACGCAATCTGCTGCTAGCTAGGTAGAGGAAGTATG 594
  
```

RESULT 2

US-09-080-983-1

Sequence 1, Application US/C9080983  
 Patent No. 6197948  
 GENERAL INFORMATION:  
 APPLICANT: Zhu, Hai-Ying  
 APPLICANT: Ling, Kai-Shu  
 APPLICANT: Gonsalves, Dennis  
 TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS  
 TITLE OF INVENTION: AND THEIR USES  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Nixon, Hartgrave, Devans & Doyle LLP

STREET: Clinton Square, P.O. Box 1051  
 CITY: Rochester  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 14603  
 COMPUTER READABLE FORM:  
 MEDIA TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/080,933  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA: US 60/047,194  
 APPLICATION NUMBER: 20-MAY-1997  
 FILING DATE: 20-MAY-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Goldman, Michael L.  
 REGISTRATION NUMBER: 30,727  
 REFERENCE/DOCKET NUMBER: 19603/1631  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (716) 263-1304  
 TELEFAX: (716) 263-1603  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 15500 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-09-080-983-1

Alignment Scores:  
 Pred. No.: 1,566-125 Length: 15500  
 Score: 991.00 Matches: 198  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-613-486-15 (1-198) x US-09-080-983-1 (1-15500)

```

QY 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20
DB 13584 ATGGAGTGTGATGCCGACGACCACTTACCACTCTGTGATTAACCGCGCTCTAGTCTA 13643
QY 21 AsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40
DB 13644 AATGGTGTCCACAGAAGCTTTTATCTGCTGAAGTTGAAAAAATGTTGGTGCAGAAAGG 13703
QY 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyrrAlaLeuAlaAATG 60
DB 13704 GCTCCTAAGCGGGTATAGAGTGTGTTGGGTCTACTCTTACGCACCTCGCGGCAAGA 13763
QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
DB 13764 ACCACGTCCTTAAAGGTTTACGCGCGCAGATTACAGAGCTTATATTTTCAAAATAGTTTCGGA 13823
QY 81 GluArgAsnValValThrGluGlyAspLysLysValIleAspGlyCysAlaPro 100
DB 13824 GAGAGGAATGTGGTAGTAAACAGAGGTGACCTTAAAGAGGTACTCGACGGGTGTGCGCT 13883
QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaVal 120
DB 13884 CTCACCTAGGTTCACTAATAAATCTTAAAGGTTTGGTGTGCTACTCTTCACTGAGGCTTACGTT 13943
QY 121 AspPheCysIleAlaTyrrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140
DB 13944 GACTTTGTATCGGTATAGCAAAATTAACCCCAACTCAACCGCGCGCGGAATTGGGG 14003
QY 141 IleProAlaGluAspSerTyrrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
  
```

```
Db 14004 ATTCCAGCTCAAGATTGCTGCTTAGCTGCGAGATTTTCTGGGTACTTTCGCCGAGGCTCTCT 14063
Qy 161 GluLeuGlnSerArgLysMetPheAlaSerMetTyxAlaLeuLysThrGluGlyGly 180
Db 14064 GAATTACAGCAAGTAGAGAGATTCGCGAGTATGTACGCTCTAAAACTCAAGGTGGA 14123
Qy 181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgArgGluValMet 198
Db 14124 GTGGTAATACACCAAGTACAGCAATCTCGGTACGCTAGGTAGGAGGAAGTTATG 14177
RESULT 3
US-09-252-991A-1531/c
; Sequence 1531, Application US/03252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1531
; LENGTH: 2034
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1531
Alignment Scores:
Pred. No.: 0.1 Length: 2034
Score: 86.00 Matches: 35
Percent Similarity: 43.04% Conservative: 33
Best Local Similarity: 22.15% Mismatches: 76
Query Match: 9.88% Indels: 14
DB: 4 Gaps: 4
US-09-613-486-15 (i-198) x US-09-252-991A-1531 (1-2034)
Qy 24 AspLysLysLeuLeuSerAaGluValGluLysMetLeuValGlnLysGlyAaProAsn: 43
Db 864 GATCAGCGCTCCCGCTTCGCGGCTCAGCGGAGTCGCCGATGAGCGCGCCAGGA 865
Qy 44 GluGlyTiedLysValValPheGlyLeuLeuTyxAlaLeuAlaAaArgThrThrSer 63
Db 804 GCGGGCATGAGACGATCGCGGGGAGCTCGTCCGCAACCATCGCGGTCCGCGGT 745
Qy 64 ProlLysValGlnArgAaAspSerAspValIlePheSerAsnSerPheGlyGluArgAsn 83
Db 744 GACCAGCTTGCCAGCGCGAGCTGGCTTGTGAGAGAACACCGCGGGGACAGGCGCCAG 585
Qy 84 ValValValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThrArg 103
Db 684 GCCGGGTAGAAAGTCGGT-----GTCGAGAGGCTGT----- 655
Qy 104 PheThrAsnLysLeuAaGlyThrPheGlyArgThrPheThrGluAlaTyxValAspPheCys 123
Db 654 ---CTGAAAGCGCTTGAGCTCGACCGCCGAGCGCTTCAGCAGGTGTTTCGCGTTC 598
Qy 124 IleAlaTyxLysHisLysLeuProGlnLeuAsnAlaAa---AlaGluLeuGlyIlePro 142
Db 597 ATACAGGGCGGTTCGCGCATTCACCGCATTCGTCGCCCGCTAGCCGAGGATAGCGGCC 538
Qy 143 AlaGluAspSerTyxLeuAaAlaAspPheLeuGlyThrCysProLysLeuSerGLeu 162
Db 537 GCCGGCTTGAATTCGTTGGCTTGGATGGCCGCCGAGTCGTCGTGCTTCGATGAC 478
Qy 163 GlnGlnSerArgLysMetPheAlaSerMetTyxAlaLeuLysThrGluGlyGly 180
Db 477 CAGGATCGCGCCCGC-----GGGGTCCGCTTCGGATAGAACACAGCGCGCGC 430
RESUL 4
US-08-565-907A-1
; Sequence 1, Application US/08565907A
; Patent No. 5814499
; GENERAL INFORMATION:
; APPLICANT: Sylvain Moineau, Barbara
; APPLICANT: J. Holler, Peter A. Vandenberg,
; APPLICANT: Ebenezer R. Vedamuthu, Jeffrey K.
; APPLICANT: Kondo
; TITLE OF INVENTION: DNA Encoding Phage
; TITLE OF INVENTION: Abortive Infection Protein
; TITLE OF INVENTION: From Lactococcus
; TITLE OF INVENTION: Lactis, and Method of Use Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 5.25 inch,
; MEDIUM TYPE: 360 Kb storage
; COMPUTER: Acer
; OPERATING SYSTEM: MS-DOS (version 4)
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/565,907A
; FILING DATE: December 1, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: Quest 4.1-152
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; TELEX: NO. 58144996
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4467
; TYPE: Nucleotide
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION: Genomic DNA
; HYPOTHETICAL: No
; ANTI-SENSE: No N/A
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Lactococcus lactis
; STRAIN:
; INDIVIDUAL ISOLATE: W1
; DEVELOPMENTAL STAGE: N/A
; HAPLOTYPE: N/A
; TISSUE TYPE: N/A
; CELL TYPE: bacterium
; CELL LINE: N/A
; ORGANELLAE: N/A
; IMMEDIATE SOURCE:
; LIBRARY: genomic
; CLONE: SMQ-20
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: phage abortive infection
; LOCATION: N/A
; IDENTIFICATION METHOD: sequencing
; OTHER INFORMATION: DNA encoding phage
```







```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13408
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13408

Alignment Scores:
Pred. No.: 0.389 Length: 1743
Score: 81.00 Matches: 30
Percent Similarity: 47.79% Conservative: 26
Best Local Similarity: 22.39% Mismatches: 50
Query Match: 8.17% Indels: 28
DB: 4 Gaps: 5

US-09-613-486-15 (1-198) x US-09-252-991A-13408 (1-1743)
QY 17 AlaSerSerLeuAsnGlyValAspLysLysLeuSerAlaGluValGlyLysMetLeu 36
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 68 GCCAACACCATGAACGGTGTCTACCGGAGGCGATGGCGAAGACCATCGCAGCGTGGAG 127
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 37 ValGlnLysGlyAlaProAsnGluGlyIleGluValValPheGlyLeuLeuTyrAla 56
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 128 GCGGAGAAA-----GAGGGGATCGCC-----GGCTGGTGGTGC 160
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 57 LeuAlaAlaArgThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSer 76
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 161 -----ACCTCGCGAAGAAAGACCTTCTCGCGGAGGCGGATCTC----- 199
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 77 AsnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLysValLeuAsp 96
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 200 -----AACGAGCTGATCAAGGTCACCAAGCGCGGCGGCGCCCGCCCTTCTACCGAG 247
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 97 GlyCysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr 116
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 248 GGCATCTCGAAGTCAAGGCGCAACTCGCGGCGCTGGAGACCTCGGCAAGCGGTGGTC 307
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 117 GluAlaTyr-----ValAspPheCysIleAlaTyrLysHis 128
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 308 GCGCGCATCAACGGCGCTCGCTGGCGCGGCGTGGAGATCTGCTGGCTGCCACAC 367
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 129 LysLeuProGlnLeuAsnAlaAlaGluLeuGlyLeuPro 142
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 368 CGCATCGCCCTGGACATCCCGCGCTGGAGTCCGCTCGGCTGCCG 459
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
US-09-252-991A-13503
; Sequence 13503, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13503
; LENGTH: 2139
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13503

Alignment Scores:
Pred. No.: 0.651 Length: 2139
Score: 81.00 Matches: 30
Percent Similarity: 47.79% Conservative: 26
Best Local Similarity: 22.39% Mismatches: 50
Query Match: 8.17% Indels: 28
DB: 4 Gaps: 5

US-09-613-486-15 (1-198) x US-09-252-991A-13503 (1-2139)
QY 17 AlaSerSerLeuAsnGlyValAspLysLysLeuSerAlaGluValGlyLysMetLeu 36
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2328 GCCAACACCATGAACGGTGTCTACCGGAGGCGATGGCGAAGACCATCGCAGCGTGGAG 2269
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
US-09-252-991A-13950/c
; Sequence 13950, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13950
; LENGTH: 2433
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13950

Alignment Scores:
Pred. No.: 0.651 Length: 2433
Score: 81.00 Matches: 30
Percent Similarity: 47.79% Conservative: 26
Best Local Similarity: 22.39% Mismatches: 50
Query Match: 8.17% Indels: 28
DB: 4 Gaps: 5

US-09-613-486-15 (1-198) x US-09-252-991A-13950 (1-2433)
QY 17 AlaSerSerLeuAsnGlyValAspLysLysLeuSerAlaGluValGlyLysMetLeu 36
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2328 GCCAACACCATGAACGGTGTCTACCGGAGGCGATGGCGAAGACCATCGCAGCGTGGAG 2269
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

QY 37 ValGlnLysGlyAlaProAsnGluGlyLeuValValPheGlyLeuLeuLeuTyraAla 56  
 Db 2268 GCGGAGAAA-----GAGGGGATCGCC-----GGGTGGTGTG----- 2236  
 QY 57 LeuAAlaArgThrSerProLysValGlnArgAlaAspSerAspValIlePheSer 76  
 Db 2235 -----ACCTCGCGAAGAAGACCTTCTTCGCGCGCGGCGATCTC----- 2197  
 QY 77 AsnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLysValLeuAsp 96  
 Db 2196 -----AAGAGCTGATCAAGGTCACCAAGCGCGAGCCCGCCGCTTACCAG 2149  
 QY 97 GlyCysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr 116  
 Db 2148 GGCATCTCGAATCAAGGGCAACTGCGCGCGCTGGAGACCTCGGCAAGCGGTGCTC 2089  
 QY 117 GluAlaTyR-----ValAspPheCysIleAlaTyRlyshis 128  
 Db 2088 GCGCGGATCAACGGCGCTGCGTGGCGCGCGCTGGAGATCTGCTGCGCTGCCACCAC 2029  
 QY 129 LysLeuProGlnLeuAsnAlaAlaGluLeuGlyIlePro 142  
 Db 2028 CCATCGCGCTGCACATCCCGCTGCGAGCTCGCGCTGCCG 1987

## RESULT 10

US-08-770-544-15

Sequence 15, Application US/08770544

Patent No. 5907085

GENERAL INFORMATION:

APPLICANT: Gonsalves, Dennis

APPLICANT: Ling, Kai-Shu

TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND

TITLE OF INVENTION: THEIR USES

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hargrave, Devans &amp; Doyle LLP

CITY: Rochester

STATE: New York

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DCS

SOFTWARE: Patent It. Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

FILING DATE: US/08/770,544

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 6000908

FILING DATE: 21-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 19603/621

TELEPHONE: (716) 263-1304

TELEFAX: (716) 263-1600

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 1434 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-770-544-15

Alignment Scores:

Pred. No.: 0.543

Score: 79.00

Percent Similarity: 40.49%

Length: 1434

Matches: 43

Conservative: 40

Best Local Similarity: 20.98% Mismatches: 60  
 Query Match: 7.97% Indels: 62  
 Gaps: 2

US-09-613-486-15 (1-198) x US-08-770-544-15 (1-1434)

QY 12 LeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLysLeuLeuSerAlaGlu 31  
 Db 904 TTCCCATTAACGAAGCTCTACAGATAAC-----GCAAGA 939  
 QY 32 ValGluLysMetLeuValGlnLysGly-----AlaProAsnGly---GlyIleGlu 47  
 Db 940 CTACGAGACTCGCTTCTTCGAAAGCGGAGCAACACACGAGATATGGGAATATG 999  
 QY 48 ValValPheGlyLeuLeuLeuTyRAlaLeuAlaAlaArgThrThrSerProLysValGln 67  
 Db 1000 ATAGTGGCATGATACAACTTTTCGTACTCTCTCTACTGTAAAGAAATATAAGCGTCAAA 1059  
 QY 68 -----ArgAlaAspSerAspValIlePheSerAsnSerPheGlyGluArgAsnVal 84  
 Db 1060 GACGGGTATAGGTGGAGACCGAATTA-----GGTCAAGAGAGATC 1101  
 QY 85 ValValThrGluGlyAspLeuLysValLeu-----AspGlyCysAlaPro 100  
 Db 1102 TACTTAAGTTATTTCGGAAGTAAGGAAGCTATATTAGGAGGAAATACGTGCGTCTCCA 1161  
 QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyRVal 120  
 Db 1162 -----ACCAACACTGTCGATCTCTCATGAGGTATTTTCTCACCACCTAT 1209  
 QY 121 AspPheCysIleAlaTyRlyshisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140  
 Db 1210 ACTCTACTTATACAGAAGAAATTCAG---CCAGCGTGTACTGCCCTAGTAAGCACGGC 1266  
 QY 141 IlePro-----AlaGluAspSerTyRLeu 148  
 Db 1267 GTCCCGAAGAGTTCACTCGTACTGCTCGACTTCGACTCTGGATACAGATATTAC 1326  
 QY 149 AlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnSerArgLysMet 168  
 Db 1327 CCGGCGGACGTGTG-----AAGGCTAACGCA 1353  
 QY 169 PheAlaSerMetTyRAlaLeuLysThrGluGlyValValAsnThrProValSerAsn 188  
 Db 1354 ATGGCTTCGCTATAGCATTAATCA-----GCTAAT 1386  
 QY 189 LeuArgGlnLeuGly 193  
 Db 1387 TTAAGGCGTAAAGGT 1401

## RESULT 11

US-09-579-259-15

Sequence 15, Application US/09579259

Patent No. 6558953

GENERAL INFORMATION:

APPLICANT: Gonsalves, Dennis

APPLICANT: Ling, Kai-Shu

TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS

PROTEINS AND THEIR USES

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hargrave, Devans &amp; Doyle LLP

STREET: Clinton Square, P.O. Box 1051

CITY: Rochester

STATE: New York

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent It. Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

```

QY      189 JcuArgGlnLeuGly 193
      :|||||:
      :|||||:
Db      1387 TTAAGCGTAAAGGT 1401

RESULT 12
US-C9-214-808-1
; Sequence 1, Application US/09214808A
; Patent No. 6475793
; GENERAL INFORMATION:
; APPLICANT: Rosenthal, Andre
; APPLICANT: Freiberg, Christoph
; APPLICANT: Perret, Xavier Philippe
; APPLICANT: Broughton, William John
; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic
; Patent No. 6475793
; TITLE OF INVENTION: Plasmid
; FILE REFERENCE: CARP0068
; CURRENT APPLICATION NUMBER: US/09/214,808A
; CURRENT FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: PCT/IB97/00950
; PRIOR FILING DATE: 1997-07-10
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1

```

```

: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO.:
: LENGTH: 536165
: TYPE: DNA
: ORGANISM: Rhizobium
JS-09-214-808-1

Alignment Scores:
Pred. No.: 5,2e+03
Score: 536165
Percent Similarity: 51
Best Local Similarity: 24
Query Match: 65
DE: 56
      9
      4

```

```

Pred. NO.: 5.25+0.3
Score: 79.00
Length: 536.63
Match: 51
Percent Similarity: 38.27
Conservative: 24
Best Local Similarity: 36.02
Mismatch: 65
Query Match: 7.97
Indels: 56
DB: 4
Gaps: 9
US-09-614-486-15 (1-198) X US-09-214-ACB-1 (1-436165)

```

```

Percent Similarity: 58.27%      ConservAtive: 24
Best Local Similarity: 26.02%    Mismatches: 65
Query Match: 7.97%       Indels: 56
DB: 4                      Gaps: 9

US-09-613-486-15 (1-198) x US-09-214-808-1 (1-536165)

QY   1 MetGluLeuMetSerAspSerAsnLeuValIleThrAspAlaSerLeu 20
      ::|||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||::
Db   314692 ATCGAAAGCTTCGGCTTCTAGTCTCCGGTCATGATACCGGCAAGCC----- 314742

QY   21 AsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   314743 ---GGAACCGGCACAAAACCTCTGGCACGGGTATTCATGAAGCTCCAGAGCTCCGCCG 314793

QY   41 -----AlapcAsr 43
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   314800 AGCAAGCTCATCAGCTCAGAAGACGAAGTGGGAATCGCTTTTGCTTCGCCCGCAGC 314859

QY   44 GluGlyIleGluValIlePheGlyLeuLeuLeuTyzAlaLeuAalaArgThrThrSer 63
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   314860 GAGGCGCGCAGGAAGGTGCTTGGGCTACTCCAATACCACCCCAAACCCCTGCTGATC 314919

QY   64 ProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGlyGluArgAsn 83
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   314920 GAGAGCGTCGAATGCCTGTCTGCTGACGCT-----CAGGAACGCTCTG 314961

QY   84 ValValValThrGlu---GlyAspLeuLysValLeuAspGlyCysAlaProLeuThr 102
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db   314962 ATAGAAGTTCATCGAGACAGCGGCTTATCGCGGCTGGGGACAAC-----CAGAGAGGG 315015

QY   103 ArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrlValAspPhe 122
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db   315016 CGGTTGGAAGCGCGGCTG----- 315033

QY   123 CysIleAlaTyrlLysHisLysLeuProGlnLeuAsnAlaIleAlaGluLeuGlyIlePro 142
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db   315034 ATCCTGGCATCGACGCGGCGCTTCCTGAGTTGGGCTCGAGCGGGCAGCTG---ATACCT 315090

QY   143 AlaGluAspSerTyrLeuAla-----AlaAspPheLeu 153
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 8, 2003, 01:33:59 : Search time 302 Seconds  
(without alignments)  
2092.035 Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 991

Sequence: 1 MELMSDSNLSNLVITDASSL.....GGWNTFVSNLRQLGRREV 195

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2141354 seqs, 1595473879 residues

Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n model -DEV=xlh  
-Q/cgn2\_1/USPFO.spool/US09613486/runat\_07112003\_120413\_27179/app\_query.fasta.i.391  
-DB=Published Applications NA -QFM=fastap -SCFPI=nrpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -XATRX=blotum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUT=EM=ptc -NOR=ext -HEAPSIZE=500 -WINLEN=3  
-MAXLEN=2000000000 -USER=US09613486 %CGN 1 1 221 @runat\_07112003\_120413\_27179  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
c 1	83.5	8.4	640681	US-09-790-988-1

2	79	8.0	1356	9	US-09-815-242-6624	Sequence 6624, Ap
3	79	8.0	1434	12	US-10-138-842A-15	Sequence 15, Appl
4	79	8.0	1434	12	US-10-039-112-15	Sequence 15, Appl
5	79	8.0	11597	10	US-09-970-927A-222	Sequence 222, App
6	79	9.0	536165	11	US-09-939-964-1	Sequence 1, Appl
7	77.5	7.8	665	10	US-09-974-300-5760	Sequence 5760, Ap
8	77	7.8	2091	12	US-10-032-585-6808	Sequence 6808, Ap
9	76.5	7.7	1068	14	US-10-156-761-4239	Sequence 4239, Ap
10	76.5	7.7	8095	10	US-09-980-107-2228	Sequence 2228, Ap
11	76.5	7.7	9025608	14	US-10-156-761-1	Sequence 1, Appl
c 12	75	7.7	600	9	US-09-864-761-8204	Sequence 8204, Ap
13	75.5	7.6	29729	10	US-09-970-927A-238	Sequence 238, App
14	75	7.6	81940	10	US-09-759-5083-1	Sequence 1, Appl
15	75	7.6	81940	12	US-09-973-319-724	Sequence 724, App
16	75	7.6	81940	12	US-09-960-706-1092	Sequence 1092, App
17	74.5	7.5	876	14	US-10-156-761-3299	Sequence 3299, Ap
18	74.5	7.5	1320	10	US-09-974-300-1589	Sequence 1589, Ap
c 19	74.5	7.5	3684	12	US-10-133-013-109	Sequence 109, App
c 20	74.5	7.5	3309400	10	US-09-738-626-1	Sequence 1, Appl
c 21	74	7.5	1161	10	US-09-954-531-575	Sequence 575, App
22	74	7.5	1612	14	US-10-205-823-376	Sequence 376, App
23	74	7.5	1715	9	US-09-925-301-543	Sequence 543, App
24	73.5	7.4	786	10	US-09-894-844-129	Sequence 129, App
25	73.5	7.4	1353	14	US-10-156-761-4968	Sequence 4968, Ap
c 26	73.5	7.4	2350	10	US-09-970-927A-109	Sequence 109, App
27	73	7.4	1020	10	US-09-938-842A-2260	Sequence 2260, Ap
28	73	7.4	1455	14	US-10-156-761-5889	Sequence 5889, Ap
29	73	7.4	5964	12	US-10-032-585-6518	Sequence 6518, Ap
c 30	73	7.4	49136	10	US-09-768-877-1	Sequence 1, Appl
c 31	73	7.4	59065	12	US-10-135-696-3	Sequence 3, Appl
c 32	73	7.4	9025608	14	US-10-156-761-1	Sequence 1, Appl
33	72.5	7.3	551	9	US-09-976-451-1	Sequence 1, Appl
c 34	72.5	7.3	3657	9	US-09-976-451-10	Sequence 10, Appl
35	72	7.3	1071	10	US-09-974-300-2722	Sequence 2722, Ap
36	72	7.3	1084	10	US-09-974-300-1197	Sequence 1197, Ap
37	72	7.3	1084	10	US-09-974-300-1308	Sequence 1308, Ap
38	72	7.3	2709	12	US-10-087-402-5	Sequence 5, Appl
39	72	7.3	3360	12	US-10-032-585-6333	Sequence 6333, Ap
c 40	72	7.3	5715	14	US-10-128-714-2237	Sequence 2237, Ap
c 41	72	7.3	5715	14	US-10-128-714-7237	Sequence 7237, Ap
c 42	72	7.3	5819	14	US-10-128-714-1237	Sequence 1237, Ap
c 43	72	7.3	5819	14	US-10-128-714-6237	Sequence 6237, Ap
c 44	72	7.3	7425	14	US-10-114-170-212	Sequence 212, App
c 45	72	7.3	7819	14	US-10-128-714-237	Sequence 237, App

ALIGNMENTS

RESULT 1  
US-09-790-988-1/c  
; Sequence 1, Application US/C9790988  
; Patent No. US20020127687A1  
; GENERAL INFORMATION:  
; APPLICANT: SHIGENOBU, SHUJI  
; APPLICANT: WATANABE, HIJEMI  
; APPLICANT: HATTORI, MASAHIRA  
; APPLICANT: SAKAKI, YOSHIOYUKI  
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
; FILE REFERENCE: 081356/0159  
; CURRENT APPLICATION NUMBER: US/09/790,988  
; CURRENT FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: JP2000-107160  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 640681  
; TYPE: DNA  
; ORGANISM: Buchnera sp.  
US-09-790-988-1  
Alignment Scores: 3.46e+03 Length: 640681  
Pred. No.: 1

```
Score: 83.50      Matches: 32
Percent Similarity: 37.18%      Conservative: 26
Best Local Similarity: 20.51%      Mismatches: 53
Query Match: 8.43%      Indels: 45
DB: 10      Gaps: 5

US-09-613-486-15 (1-198) x US-09-790-988-1 (1-640681)

QY 65 LysValGlnArgAlaSerAspValIlePheSerAsnSerPheGlyGluArgAsnVal 84
DB 435915 AAAAAATCAACGTCGATCATTAAGCATTTGATGACGAAGGTACTCCAGCGTCAAAAGAATATT 435856

QY 85 ValValThrGluGlyAspLeuLysLysValLeuAsp----- 96
DB 435855 TTAATTGAAATCGAATATATAAAATAATATGCAAGATAAAGCTTAATCGCGGTTTAATG 435795

QY 97 -----GlyCysAlaProLeuThrArg 103
DB 435795 GGTGTTAAGTCTACCGGAAATGTCGTCTGATCTTATTCGTGCTACCTATGCTCGT 435736

QY 104 PheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrrValAspPheCys 123
DB 435735 AGCATTAAT-----ACCATATG 435718

QY 124 IleAlaTyrrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAla 143
DB 435717 TTATCTGGAAGCTCTAAATAGATGATATAATTAAGAGCGTCGACTATGGAATA----- 435664

QY 144 GlnAspSerTyrrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGln 163
DB 435663 -----TATCTCTGAAATTTTCTGGA-----GGACAGGTAGATATTACT 435625

QY 164 GlnSerArgLysYerPheAlaSerMetTyrrAlaLeuLysThrGluGlyValValAsn 183
DB 435624 TCCGGAAATTTGTATTTCTACTTCGGAGCTTATTTAATTAATAAATGGAATAATTTGT 435565

QY 184 ThrProValSerAsnLeuArgGlnLeuGlyArg-----ArgGluValMet 198
DB 435564 ACACCAATTAATAAACACTACTCTCAATAGGATCAGGATTAGAAGTTATG 435517

RESULT 2
US-09-815-242-6624
; Sequence 6624, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Christen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant C.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in:
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/197,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6624
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1356)
US-09-815-242-6624

Alignment Scores:
Pred. No.: 0.851      Length: 1356
Score: 79.00      Matches: 53
Percent Similarity: 36.53%      Conservative: 27
Best Local Similarity: 24.20%      Mismatches: 73
Query Match: 7.97%      Indels: 66
DB: 9      Gaps: 12

US-09-613-486-15 (1-198) x US-09-815-242-6624 (1-1356)

QY 1 MetGluLeuValSerAspSerAsnLeuSer-----AsnLeuValIleThrAspAla 17
DB 529 TTAGATTACGACAGAGGACTCGAATGCTGCGAGTGGATATGAACATCGTGATGACCGAGCT 588

QY 18 SerSerLeu-----AsnGlyValAspLysLysLeuSerAlaGluVal 32
DB 589 GGTCAATTTGTGGAATTCAGGAACCGGTGAGAGACTACCTTCTCAGGAGATGAGTTA 648

QY 33 GluLysMetLeuValGlnLysGlyAlaProAsnGluGlyLeGluValValPheGlyLeu 52
DB 649 AATCCCATGCTCTTTTAT-----GGAAAAACAGGATCGAGGAATTAATTCGCTAT 699

QY 53 -----LeuLeuTyrrAlaLeuAlaArgThrThrSerProLysValGlnArgAla 69
DB 700 CAGAAAGACGCTTTATGCTTTAGCC-----CTGAGGAAGTCCCTCCCAA 747

QY 70 AspSerAsp-----ValIle-----Phe 75
DB 748 GATTCAAGAGGAAGTGTATGTCATTGCAACTAGAAATCTGTGTAAGCGGAAAGACTTT 807

QY 76 SerAsnSerPheGlyGluArgAsrValValValThrGluGlyAspLeuLysValLeu 95
DB 808 AGTTCAATTTTGGTGAAGAGGATATACAGTT-----AAGACGTTATTA 852

QY 96 AspGlyCysAlaProLeuThrArgPheThrAsn-----LysLeuArgThrPheGlyArg 113
DB 853 GAT-----TATCCGAAATTGCGCAGATGTAGAGAAACTGGTCTGA 891

QY 114 ThrPheThrGluAlaTyrrValAspPheCysIleAlaTyrrLysHisLysLeuProGlnLeu 133
DB 892 ACATTTGAAGAA-----AATGCTCGCTAAAGCTGAACCAATT 930

QY 134 AsnAlaAlaGluLeuGlyIleProAlaGluAspSerTyrrLeuAlaAlaAspPheLeu 153
DB 931 GCAGAAATTTTACAGAAACCAGTATTAGCAGATGATTTCAGGATTAATTTGTTGATGCCGTTA 990

QY 154 GlyThrCysProLysLeuSerGluLeuGlnSerArgLysMetPheAlaSerMetTyr 173
DB 991 GGGGAATGCCG-----GGATCTAT 1011

QY 174 AlaLeuLysThrGluGlyGlyValValAlaAsnThrProValSerAsnLeuArgGlnLeu 192
DB 1012 TCTGCTGTTTTGCTGCTGAGCCAACTAACGATGCTTCAAAATAATGCGAAACTACTA 1068

RESULT 3
US-10-138-842A-15
; Sequence 15, Application US/10138842A
; Publication No. US20030148390A1
; GENERAL INFORMATION:
; APPLICANT: GONSALVES, DENNIS
; APPLICANT: LING, KAI-SHU
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND
; TITLE OF INVENTION: THEIR USES
```

RESULT 4  
US-16-039-112-15  
Sequence 15, Application US/10039112  
Publication No. US20030198942A1  
GENERAL INFORMATION:  
APPLICANT: Gonsalves, Dennis  
          Ling, Kai-Shu  
TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS  
                      PROTEINS AND THEIR USES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
STREET: Clinton Square, P.O. Box 105,  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14603  
COMPUTER READABLE FORM:

```
,  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOCS  
SOFTWARE: Patent In-Re-case #1.0, Version #1.30  
  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/039,112  
FILING DATE: 31-Dec-2001  
CLASSIFICATION: <Unknown>
```

US-09-613-486-15 (1-198) x US-10-138-842A-15 (1-1434)

ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 263-1304  
TELEFAX: (716) 263-1600

SEQUENCE CHARACTERISTICS:  
LENGTH: 1434 base pairs  
TYPE: nucleic acid

```

;
; CIRCULARNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:

```

03-10-033-112-13	
Alignment Scores:	
Pred. No.:	0.934
Length:	1434

Score:	79.00	Matches:	40
Percent Similarity:	40.49%	Conservative:	40
Best Local Similarity:	20.98%	Mismatches:	60
Query Match:	7.97%	Indels:	62

US: 12 Gaps.  
US-09-613-486-15 (1-198) x US-10-039-112-15 (1-1434)

924 TTGCCGATACCGAGCTCTACAGATAAAC-----  
 12 leuvalle..llrAspParaseiserleuAsnInG..yValAspLyslys  
 QY ||||| |||||:::|||||:::|||||

32 valglnlysmetleuvalgmllysgly-----atpfloaslnl  
 940 ctacgcagactcgttcttttcgaaggggggagtcacacaccacgag

Cy  
48 va-ai-phesiyLeuLeuLeuLyfAlaLeuAlaAA-aafg-nfinr :  
::: ||| :: |::: ||| :::  
Db 1000 ATAGTGGCAATGATACAACTTTTCGTACTCTACTCTACTGTAAGA

```

Qy 68 -----ArgAlaAspSerAspValIlePheSerAsnSerPhe
      |||  :::::
Qb 1060 GACGGGTATAGGTGGAGACCGAATTA-----G

```

85 ValValThrGluGlyAspLeuLysLysValLeu-----A  
Qy





```

RESULT: 7
US-09-974-300-576C
Sequence 5760, Application US/39974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berkla, Randy M.
APPLICANT: Clausen, Ib Grcth
TITLE OF INVENTION: Methods For Monitor
TITLE OF INVENTION: Expression
FILE REFERENCE: 10085.500-US
CURRENT APPLICATION NUMBER: US/09/974
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/680.598
PRIOR FILING DATE: 2003-10-06
PRIOR APPLICATION NUMBER: 60/279.526

```

FILE REFERENCE: 10182-005-999  
CURRENT APPLICATION NUMBER: US/10/032.585  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 8000  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 6808  
LENGTH: 2091  
TYPE: DNA  
ORGANISM: Candida albicans  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1209)..(1209)  
OTHER INFORMATION: n-g, a, t o r c  
US-10-032-585-6808

Alignment Scores:  
Pred. No.: 3.3 Length: 2091  
Score: 77.00 Matches: 39  
Percent Similarity: 35.68% Conservative: 32  
Best Local Similarity: 19.60% Mismatches: 74  
Query Match: 7.77% Indels: 54  
DB: 12 Gaps: 8

US-09-613-486-15 (1-198) x US-10-032-585-6808 (1-2091)

QY 2 GluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsn 21  
|||||  
Db 703 GAATATTGAAACACACAAACCAATCA-----TATTCGATAAATTACAGAAATTTGGAA 756  
|||||  
QY 22 GlyValAspLysLeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAla 41  
|||||  
Db 757 AATAATGAATCCAAATATTACAAATTCGAAATTCGAGAAGCTT----- 799  
|||||  
QY 42 ProAsnGluGlyLeuGluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAaArgThr 61  
|||||  
Db 799 -----CAATAGAAACATAAATATTAATGATTGTTTAAAGGCAATG----- 837  
|||||  
QY 62 ThrSerProLysValGlnAlaAlaAspSerAspValIlePheSerAsnSerPheGlyGlu 81  
|||||  
Db 838 -----GATACAGCTATTGCTCTCATCAATGAATATTAAGAT 873  
|||||  
QY 82 Arg-----AsnValValValThrGluGlyAspLeuLysLysValLeuAspGlyCys 98  
|||||  
Db 874 AACGAGAACACCGATGATACACCAATCATACCAAGTAAGTAAAGATTTTATGATATTTGT 933  
|||||  
QY 99 AlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAla 118  
|||||  
Db 934 -----AATATCTTCAAGCTAAATACTTACTTTTACAGAGAAA 972  
|||||  
QY 119 TyrValAspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaGlu 138  
|||||  
Db 973 TATGACTCTAAATAATACAGGTGAAGAATTAATAAGAAATAGAGATCAGCTAGA 1032  
|||||  
QY 139 LeuGlyLeuProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLys 158  
|||||  
Db 1033 -----GAAATGAACAGATTTTCTT-----CCTAACG 1059  
|||||  
QY 159 LeuSerGluLeuGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGlu 178  
|||||  
Db 1060 AATACGTATTAGAATCTAAATTAAGATATATATGCTGACCAA----- 1101  
|||||  
QY 179 GlyGlyValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgArgGluVal 197  
|||||  
Db 1102 -----AATACCAAAATTTGAACACAAAATCATTACGTGACAGGAATTT 1146  
|||||

RESULT 9

US-10-156-761-4239  
Sequence 4239, Application: US/10156761  
Publication No. US20030119018A1  
GENERAL INFORMATION:  
APPLICANT: OMURA, SATOSHI  
APPLICANT: IKEDA, HARUO  
APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIYUKI  
APPLICANT: HATTORI, MASAHIRA  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-262  
CURRENT APPLICATION NUMBER: US/10/156.761  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: JP 2001-204089  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 4239  
LENGTH: 1068  
TYPE: DNA  
ORGANISM: Streptomyces avermitilis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1068)  
US-10-156-761-4239

Alignment Scores:  
Pred. No.: 1.34 Length: 1068  
Score: 76.50 Matches: 34  
Percent Similarity: 40.71% Conservative: 12  
Best Local Similarity: 30.09% Mismatches: 38  
Query Match: 7.72% Indels: 29  
DB: 14 Gaps: 6

US-09-613-486-15 (1-198) x US-10-156-761-4239 (1-1068)

QY 57 LeuAlaAlaArgThrThrSerProLysValGlnArgAlaAspSerValIlePheSer 76  
|||||  
Db 715 CTCGCGCGGCCACCGGAGGCTGGGTACCAACCGCGACCCGCGG---CTGTACGAG 771  
|||||  
QY 77 AsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLysLysValLeuAsp 96  
|||||  
Db 772 AACGCAACCCCTGAACATCGGTTCAGGCCATTCGCGGACAGCGCGAGAGCTGGCCGAC 831  
|||||  
QY 97 GlyCysAlaProLeuThrArgPheThrAsnLysLeu---ArgThrPheGlyArgThrPhe 115  
|||||  
Db 832 CGGTGCGCGCGATCGCGCGCGCTGAGTCCCTCGACGCGTTCGTGACATTC---GGC 945  
|||||  
QY 116 Thr-----GluAlaTyrValAspPheCysIleAla 125  
|||||  
Db 892 ACCCGGACCGCGCTCGCGCGCTGAGTCCCTCGACGCGTTCGTGACATTC---GGC 945  
|||||  
QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIle----- 141  
|||||  
Db 946 GCGCGCCAC-----ATGAGCTGGGCGATCACCAGATCGCG 981  
|||||  
QY 142 -----ProAlaGluAspSerTyrLeuAlaAlaAsp 151  
|||||  
Db 982 ATCCACTGGCGCCATCCCGGCTCGGACTTCGCGCGGAC 1020  
|||||

RESULT 10

US-09-883-107-2228  
Sequence 2228, Application: US/09880107  
Patent No. US20020142981A1  
GENERAL INFORMATION:  
APPLICANT: Horne, Darcie T.  
APPLICANT: Vockley, Joseph G.  
APPLICANT: Scherf, Uwe  
APPLICANT: Gene Logic, Inc.  
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
FILE REFERENCE: 44921-5028-WO  
CURRENT APPLICATION NUMBER: US/09/880.107  
CURRENT FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: US 60/211,379  
PRIOR FILING DATE: 2000-06-14  
PRIOR APPLICATION NUMBER: US 60/237,054  
PRIOR FILING DATE: 2000-10-02

```
/ NUMBER OF SEQ ID NOS: 3950
/ SOFTWARE: Patent in Ver. 2.1
/ SEQ ID NO 2228
/ LENGTH: 8095
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURES:
/ OTHER INFORMATION: Genbank Accession No. US2002014298:Al 576927
US-09-880-107-2228

Alignment Scores:
Pred. No.: 33.7 Length: 8095
Score: 76.50 Matches: 32
Percent Similarity: 38.06% Conservatives: 19
Best Local Similarity: 73.88% X-smatches: 61
Query Match: 7.72% Indels: 22
DB: 10 Gaps: 4

US-09-613-486-15 (1-198) x US-09-880-107-2228 (1-8095)

QY 56 AlaLeuAlaAlaArgThrThrSerProLysValGlnArgAlaAspSerAspValIlePhe 75
DB 6984 GCAGTGAGAGCGGCTCTGACTCTCTTCTCTCCCGCAGCTGCCAGGAGCTGTGAGC 7043

QY 76 SerAsnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLysLysValLeu 95
DB 7044 AAGAGAGCGCTTCGGCGGGCGCGCAGCTGGTG-----GGGGAGATTCGGCGCAGCGCC 7097

QY 96 AspGlyCysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPhe 115
DB 7098 CAGCGACGCGCGCGCTCAGACGT-----GGCGACTACAGAGCTTTGGCGCGCTCATG 7151

QY 116 ThrGluAlaTyrValAspPheCysIleAlaTyrLysLysLys----- 129
DB 7152 GTGGAGAGC-----CACCGCTCACTCAGCTGAGCGCCT 7184

QY 130 -----LeuProGlnLeu-AsnAlaAlaAlaGluLeuGlyIleProAlaGluAspSerTy 147
DB 7185 CTGGCGCGCGCGCTCTCGCGGGACAGCGCGCCAGCGCCACCCCTTCAATATCTCT 7244

QY 147 rLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSerArgLy 167
DB 7245 CTGCAGAGACGACTATAGGTGAGTGTGCCAGCTGGACAGCTGGTGGAGGCTGCGCT 7304

QY 167 sMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGly 180
DB 7305 TGCTGTGCTGGGGTTATGGCAGCGCCGATGACGGCGGT 7344

RESULT 11
US-10-156-761-1
/ Sequence 1, Application US/10156761
/ Publication No. US20030119018A1
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIBA, TADAYOSH
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10156,761
/ CURRENT FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 1
/ LENGTH: 9025608
/ TYPE: DNA
/ ORGANISM: Streptomyces avermitilis

/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (4187715)
/ OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Alignment Scores:
Pred. No.: 1.84e+06 Length: 9025608
Score: 76.50 Matches: 34
Percent Similarity: 40.71% Conservatives: 12
Best Local Similarity: 30.09% Mismatches: 38
Query Match: 7.72% Indels: 29
DB: 14 Gaps: 6

US-09-613-486-15 (1-198) x US-10-156-761-1 (1-9025608)

QY 57 LeuAlaAlaAlaArgThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSer 76
DB 5216859 CTGGCGCGCGCCACGCGCAGCTGGGTGACACCGCGGACCGCGCGG---CTGTACGAG 5216975

QY 77 AsnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLysLysValLeuAsp 96
DB 5216916 AACGCGACCCCTGAACAGTCGTTCAAGCCATTCCGCGGACAGCGCGAGAGCTGCCGAC 5216975

QY 97 GlyCysAlaProLeuThrArgPheThrAsnLysLeu---ArgThrPheGlyArgThrPhe 115
DB 5216976 GGTGTGCGCGCGATCGCGCGCAGCTGAGGAGCTGGACAGGTCTCTCACCAGCGCTTC 5217035

QY 116 Thr-----GluAlaTyrValAspPheCysIleAla 125
DB 5217036 ACCCGGACCGCGGTGCGCGCTGAGTCCCTCGACGCGTTCGTGCTGACTTC-----GCG 5217089

QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaAlaGluLeuGlyIle----- 141
DB 5217090 GGCGCGCCAC-----ATGGAGCTGGGGCATCACCGACATCGCG 5217125

QY 142 -----PrcAlaGluAspSerTyrLeuAlaAlaAsp 151
DB 5217126 ATCCACTGGCCCATCCCGACTCCGACTTCGCGCGCGGAC 5217164

RESULT 12
US-09-864-761-8204/c
/ Sequence 8204, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Pean, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aestica-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/006666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
```

2



US-09-613-486-15 (1-198) x US-09-873-319-724 (1-8-940)

```
QY      54 LeuTyrAlaLeuAlaAlaArgThrThrSer-----ProTys 65
Db      64723 GTATACAGGTGGAGCCAAATAATGCTGGTCTGCATAAGCAGAAATTAAGTGAATA 64782
QY      66 ValGlnArgAlaAspSerAspValIle-----PheSerIasnSerPheGlyGlu 81
Db      64783 GTACAGATACACCAGGAAGTAGTTGGGCCAATAAGATTCCCAATATTACTGGGGAG 64842
QY      82 ArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaProLeu 101
Db      64843 AAGATGACTCTGTGGTGGATGCCCCACTCAAT-----GACGGTGTGCTCCCATTA 64893
QY      102 ThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAsp 121
Db      64894 ACCCACTACA-CATTGAANAACCGGMAACCCAGCAGACTTGCTGGGCACCTAATTGAGGAT 64953
QY      122 PheCysIleAla-----TyrLysHis 128
Db      64954 AAATGTGAAGCCCAAGTTACACTGCCATTAACTATAAACGGCAATGAATACCAATTTC 65013
QY      129 LysLeuProGlnLeuAsnAlaAlaAlaGluLeuGlyIleProAlaGluAspSerTyrLeu 148
Db      65014 CGTGTTTCTGCAGTTAACAAAGTTTGGT---GTTGGCAGGCCACTTGATTCTGATCCAGTG 65073
QY      149 AlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSerArgLysMet 166
Db      65073 GTTGCTCAATACATATATCTGTTCTGATGCCCTGGCAATCCAGAACCTAGCAACATA 65130
QY      169 PheAlaSerMetTyrAlaLeu-----LysThrGluGlyGlyValValAsn 193
Db      65131 ACAGCAACACGCAATTACCCCTGCATCGGCAAGGCCAGCAATCAGATCGTGGC----- 65181
QY      184 ThrProValSerAsnLeuArgGln-----LeuGlyArgArgGlu 196
Db      65182 -----ACTGAATTCACACGTATATCTCTTGAAGAAGAGAA 65217
```

Search completed: November 8, 2003, 03:17:34  
Job time : 2632 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.  
November 7, 2003, 14:47:08 ; Search time 21 Seconds  
(without alignments)  
906.733 Million cell updates/sec  
US-09-613-486-15  
Perfect score: 991  
Sequence: 1 MELXSDNSLSNVITDASSL.....GWNTPVSNLRQLGERVM 198  
Scoring table: BL0SUM62  
Gapop 16.0 , Gapext 0.5

Searched: 281308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 281308  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	293	29.6	204	2 S27900	coat protein - sug
2	286	28.9	204	2 S28715	coat protein - sug
3	125	12.6	223	2 D49804	capsid protein p25
4	104	10.5	216	2 S27899	hypothetical prote
5	100.5	10.1	216	2 S34205	coat protein homol
6	100.5	10.1	240	2 C49804	capsid protein hom
7	99.5	10.0	216	2 S28714	hypothetical prote
8	92	9.3	1692	2 A33988	adenylate cyclase
9	90	9.1	419	2 A33246	P-450 monooxygena
10	87	8.8	527	2 P69673	phosphoenolpyruvat
11	86.5	8.7	328	2 B84722	probable GDSJ-mot
12	86	8.7	360	2 A84722	probable GDSJ-mot
13	86	8.7	445	2 T34352	hypothetical prote
14	86	8.7	447	2 T34352	probable beta-lact
15	85.5	8.6	729	2 T08639	hypothetical prote
16	85	8.6	1013	2 A69226	type I restriction
17	84	8.5	784	2 H87398	transketolase faml
18	83.5	8.4	483	2 E84976	tidd protein .impo
19	83	8.4	530	2 E90353	ABC transporter, A
20	83	8.4	530	2 B90370	ABC transporter SSO
21	83	8.4	56	2 AD2581	phosphoenolpyruvat
22	83	8.4	561	2 B97363	phosphoenolpyruvat
23	81.5	8.2	823	1 D70176	endopeptidase Ia (
24	81	8.2	299	2 H72217	purine nucleoside
25	81	8.2	317	2 E82815	ribokinase XF0366
26	81	8.2	714	2 F83427	probable 3-hydroxy
27	80.5	8.1	387	2 D72335	flagellin - Cherno
28	80.5	8.1	673	2 B75514	penicillin-binding
29	80	8.1	555	2 G86741	formate-tetrahydro

30	75.5	8.0	320	2 A13234	hypothetical prote
31	79	8.0	308	2 S76003	hypothetical prote
32	79	8.0	422	1 A32306	cytochrome P450 Cy
33	78.5	7.9	209	2 D86758	orotate phosphorib
34	77.5	7.8	392	2 C83151	probable two-compo
35	77.5	7.8	523	2 S75615	hypothetical prote
36	77	7.8	289	2 D71235	hypothetical prote
37	77	7.8	397	2 B71078	probable NADH oxid
38	77	7.8	538	2 A83018	probable sodium/hy
39	76.5	7.7	503	2 AB2040	hypothetical prote
40	76.5	7.7	2712	2 TC5113	hypothetical prote
41	76	7.7	283	2 TC9386	(S)-2-hydroxy-acid
42	76	7.7	384	2 TC2968	cyclin A-type (clo
43	76	7.7	835	2 T39106	vacuolar protein s
44	76	7.7	1016	2 D86308	translation initia
45	75.5	7.6	291	2 AH1680	methyltransferases

ALIGNMENTS

RESULT 1  
S27900  
coat protein - sugar beet yellows virus  
C:Species: sugar beet yellows virus, SBV  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 11-Jan-2003  
C:Accession: S27900; S34206  
R:Brundstedt, J.; Moseley, J.; Hull, R.  
submitted to the EMBL Data Library, August 1991  
A:Description: Nucleotide sequence of cDNA encoding the coat protein of beet yellows vi  
A:Reference number: S27898  
A:Accession: S27900  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-204 <BRU>  
A:Cross-references: EMBL:MS9452; NID:g323237; PID:AAA72955.1; PID:g323240  
R:Aganovsky, A.A.; Koonin, E.Y.; Boyko, V.P.; Maiss, E.; Lunina, N.A.; Atabekov, J.G.  
submitted to the EMBL Data Library, June 1993  
A:Description: Complete nucleotide sequence of the Ukrainian isolate of Beet yellows vi  
A:Reference number: S34202  
A:Accession: S34206  
A:Status: preliminary  
A:Molecule type: genomic RNA  
A:Residues: 1-204 <AGR>  
A:Cross-references: EMBL:X73475; NID:g313689; PIDN:CAA51861.1; PID:g313694  
C:Superfamily: SBV probable coat protein

Query Match		29.6%	Score 293;	DB 2;	Length 204;
Best Local Similarity		37.6%	Pred. No. 6.6e-20;		
Matches		71;	Conservative	31;	Mismatches 85;
				Indels	2;
				Gaps	2;
QY	7	SNLSNLTITDASSLNGVDKLLSAEVEKMLVQKGAPNEGIEVVVFGLLLYALAAARTTS	SPKV	66	
DB	11	ATFENVSLADOTCHGDCCKLRNFECKLKGVPDKLGLALGCLYSCATIGTSNKV	70		
QY	67	GRASDVIFNSNF-GERNVVVTEGLKKVLDGCGAPLTRF-NKLTFRGTFTTEAYVDFCIA	125		
DB	71	SVQPTSTFIKASFGGKELFLTHGELRSFLDSQKLLGKPKNKLRCFCRTFKQDYISPAKE	130		
QY	126	VKHLPLQNLAAELGIPAESYLAADFLGTCPKLSELCQSKKMFASMYAKTEGGVNTF	185		
DB	131	YRGLPPIARNRHGLPAEDHYLAADFISTSTELTDIQOQGRLLARENATHFTFS-SESP	189		
QY	186	VSNLRQLGR	194		
DB	190	VTSKQLGR	198		

RESULT 2  
S28715  
coat protein - sugar beet yellows virus  
C:Species: sugar beet yellows virus, SBV  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 11-Jan-2000



```
C:Accession: S28715
R:Agarovsky, A.A.; Boyko, V.P.; Karasev, A.V.; Lunina, N.A.; Koonin, E.V.; Dolja, V.V.
J. Gen. Virol. 72, 15-23, 1991
A:Title: Nucleotide sequence of the 3'-terminal half of beet yellows closterovirus RNA
A:Reference number: S28710; MUID:91116305; PMID:1990061
A:Accession: S28715
A:Molecule type: DNA
A:Residues: 1-204 <BR>
A:Cross-references: EMBL:X53462; NID:958878; PIDN:CAA37554.1; PID:958883
C:Superfamily: SBV probable coat protein

Query Match      28.9%; Score 286; DB 2; Length 204;
Best Local Similarity 37.6%; Pred. No. 3e-19;
Matches 71; Conservative 29; Mismatches 87; Indels 2; Gaps 2;

QY 7 SNLSNLVITDASSNGVDKLLSAEVEKMLVQKGNIEGVFGLJLYALAAARTTSRKV 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 ATPENVSLADQTCLEHGEDCDKLRKNFEGLKLGVPEDNLGALGLCLYSCATIGTSNKV 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 67 QRADSDVIFSNFSGERNVVVTGDLKKVLDGCAPLTRFTNKLRTFGRTTEAYVDFCIA 125
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 NVOPSTFTFKASGGGKELVLTGELNSFGSQKLEGGPKNLKRCFRTFQDYSLRKE 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 126 YKHKLQPLNAAELGPAEDSYLAADFLGTCPKLSELOQSRKMFASMYALKTEGGVWVTP 185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 131 YRGKLPPIARANRHGLPAEDHYLAADFSTSTELTDLQOSRLLARENATHTEFS-SESP 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 186 VSNLRQLGR 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 VTSKQLGR 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
C49804
capsid protein p25 - citrus tristeza closterovirus CTV
C:Species: citrus tristeza closterovirus CTV
C:Date: 11-Oct-1994 #sequence_revision 25-Apr-1997 #text_change 11-Jan-2000
C:Accession: D49804
R:Pappu, H.R.; Karasev, A.V.; Anderson, E.J.; Pappu, S.S.; Hilf, M.E.; Febres, V.J.; Eck
S.M.; Dawson, W.O.; Lee, R.F.; Niblett, C.L.
Virology 199, 35-46, 1994
A:Title: Nucleotide sequence and organization of eight 3' open reading frames of the cit
A:Reference number: A49804; MUID:94160579; PMID:8116253
A:Accession: D49804
A>Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-223 <PAP>
A:Note: sequence extracted from NCBI backbone (NCBIN:144092, NCBI:P:144102)
C:Superfamily: SBV probable coat protein

Query Match      12.6%; Score 125; DB 2; Length 223;
Best Local Similarity 25.9%; Pred. No. 0.00036;
Matches 56; Conservative 33; Mismatches 63; Indels 64; Gaps 10;

QY 18 SSLLN-GVDKLLLSA-EVEKMLVQKGA-----PNEGIEVWFGLLYALA 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 29 SSVNLHIDPTLITMNDVRLSTQNAALNRDPLTLKKGKHPNLPBKDKDFHIAVMYRLA 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 ARTTS-----PKVQPAD---SDVIP-SNSFGERNVVVTGDLKKVLDGC 98
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 VKSSSLQSDDDATGITYTREGVEVLSDKLWTDVFNFSKIGNR----- 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 99 APLTRFTNKLRTFGRTTEAYVDFCIAKHKLQPLNAAELGIPAEEDSYLAADFLGTCPK 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 -----TNALRVWGRINDALYLAFC-RQNRNLSYGRPLDAGIPAGHYLCADFUTGAG 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 159 LSELQSRKQKFAFASMYALKTEGGVWVTPVSNLRQLGR 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 185 LTDLCAVYIAKEQLLK-KRGADVVVWTVNRQLGK 219
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
C49804
```

## S27899

Hypothetical protein 2 - sugar beet yellows virus  
C:Species: sugar beet yellows virus, SBV  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999  
A:Accession: S27899  
R:Brustedt, J.; Moseley, J.; Hull, R.  
submitted to the EMBL Data Library, August 1991  
A:Description: Nucleotide sequence of cDNA encoding the coat protein of beet yellows v  
A:Reference number: S27898  
A:Accession: S27899  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-216 <BR>  
A:Cross-references: EMBL:M59452; NID:9332337; PIDN:AAA72954.1; PID:9332339  
C:Superfamily: SBV probable coat protein

Query Match 10.5%; Score 104; DB 2; Length 216;  
Best Local Similarity 29.5%; Pred. No. 0.032;  
Matches 41; Conservative 19; Mismatches 62; Indels 22; Gaps 6;

QY 61 TTPKQVORADSDVIFNSFGERNVVVTGDLKKVLDGCAPLTRFTNKLRTFGRTTEAYV 120  
 :  
Db 83 TTSTKVNIVGA---YEYTIIGKKFLVKDAWVFPLKECMKKFNKPNVPTFCATFEDAY- 138  
 :  
QY 121 DFCIAYKHKLQPLN---AAALGCPAEDSYLAADFL---GTC-----PKLSELQOSRKMFA 170  
 :  
Db 139 ---IVIARSLPKLNLNRITGKRGIPSGYEFGLGADFLTATSVCLNDHEKAIVLQASRAAID 195  
 :  
QY 171 SMYALKTEGGVWVTPVSNLRQLGR 194  
 :  
Db 196 RAVSSSVVDGKIV-----SLPDLGR 214  
 :

## RESULT 5

S34235  
coat protein homolog - sugar beet yellows virus  
C:Species: sugar beet yellows virus, SBV  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 10-Dec-1999  
C:Accession: S34235  
R:Agarovsky, A.A.; Koonin, E.V.; Boyko, V.P.; Maiss, E.; Lunina, N.A.; Atabekov, J.G.  
submitted to the EMBL Data Library, June 1993  
A:Description: Complete nucleotide sequence of the Ukrainian isolate of Beet yellows v  
A:Reference number: S34202  
A:Accession: S34205  
A>Status: preliminary  
A:Molecule type: genomic RNA  
A:Residues: 1-216 <AGR>  
A:Cross-references: EMBL:X73475; NID:9332689; PIDN:CAAS1860.1; PID:9332693  
C:Superfamily: SBV probable coat protein

Query Match 10.1%; Score 100.5; DB 2; Length 216;  
Best Local Similarity 25.0%; Pred. No. 0.067;  
Matches 50; Conservative 27; Mismatches 94; Indels 29; Gaps 7;

QY 5 SDSNLSNLVITDASSNGVDKLLSAEVEKMLVQKGNIEGVFGLJLYALAAARTTSP 64  
 :  
Db 34 SEVNPKNLNRKEDELJGVIRERFKSELV-----ITDEDFVKHLAFALIRAAITST 86  
 :  
QY 65 KYQPADSDVIFNSFGERNVVVTGDLKKVLDGCAPLTRFTNKLRTFGRTTEAYVDFCI 124  
 :  
Db 87 KVNIVGA---YEYTIIGKKFLVKDAWVFPLKECMKKFNKPNVPTFCATFEDAY----I 139  
 :  
QY 125 AYKHKLQPL---NAAALGIPAEEDSYLAADFL---GTC-----PKLSELQOSRKMFA 174  
 :  
Db 140 VIARSLPKLFLNRTIGKRGIPSGYEFGLGADFLTATSVCLNDHEKAIVLQASRAAIDRAVS 199  
 :  
QY 175 LKTEGGVWVTPVSNLRQLGR 194  
 :  
Db 200 SSVVDGKIV-----SLPDLGR 214  
 :

## RESULT 6

C49804

capid protein homolog p27 - citrus tristeza closterovirus CTV  
C:Species: citrus tristeza C:steroovirus CTV  
C:Date: 11-Oct-1994 #sequence\_revision 25-Apr-1997 #text\_change 11-Jan-2003  
C:Accession: C49804  
R:Pappu, H.R.; Karasev, A.V.; Anderson, E.J.; Pappu, S.S.; Half, M.E.; Febres, V.J.; Eck  
S.M.; Dawson, W.O.; Lee, R.F.; Niblett, C.L.  
Virology 199, 35-46, 1994  
A:Title: Nucleotide sequence and organization of eight 3' open reading frames of the cit  
A:Reference number: A49804; MUID:94160579; PMID:8116253  
A:Accession: C49804  
A:Status: preliminary  
A:Molecule type: genomic RNA  
A:Residues: 1-240 <PAP>  
A:Cross-references: GB:U16304; GB:U02547; GB:U20760; N:D:8306738; PIDN:AA59629.1; PID:9  
A:Note: sequence extracted from NCBI backbone (NCBI:144092, NCBI:P144099)  
C:Superfamily: SBV probable coat protein

Query Match 10.1%; Score 100.5; DB 2; Length 240;  
Best Local Similarity 26.4%; Pred. No. 0.077; Mismatches 70; Indels 61; Gaps 13;  
Matches 57; Conservative 28

QY 2 ELYSSNLSNVIT-DASSLNGVCKLLSAEVEKMLVQKAPNEGIEVFGLLYALAAAR 60  
DB 61 EKFTGSH-KYVNVMTDTP-LENYKTKTELLVHLMICK-----RLVTI--- 104  
QY 61 TTSVQVQADSDVIFNSFGERNVVTGGDLK-KVLDGCA---PLTRFT-----NKRTF 111  
DB 105 STSTTKTKFKDKGC-----SYVOGGURKYLDRKVPFFIISKFTDRETNALRKF 154  
QY 112 GRTFTAYYDFCIA-----YKHKLQPLNAAALG:PAEDSYLAADF-LGCPKLSF--- 161  
DB 155 ACTFEELHJ--CMARLRPLDYENK-----RTRAGTPHLKLGSADFLSGSLPGYSEHER 207  
QY 162 ----LQSRKMPFASMLKTEGGVNTVPSNLRQLQR 194  
DB 208 GILRASSENLA-----RRQGYEATELNLNRDLCK 238

RESULT 7  
S28714  
Hypothetical protein 5 - sugar beet yellows virus  
C:Species: sugar beet yellows virus, SBV  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 10-Dec-1999  
C:Accession: S28714  
R:Aganovsky, A.A.; Boyko, V.P.; Karasev, A.V.; Lunina, N.A.; Koonin, E.V.; Dolja, V.V.  
J. Gen. Virol. 72, 15-23, 1991  
A:Title: Nucleotide sequence of the 3'-terminal half of beet yellows C:steroovirus RNA  
A:Reference number: S28710; MUID:91116305; PMID:1990061  
A:Accession: S28714  
A:Molecule type: DNA  
A:Residues: 1-216 <AGR>  
A:Cross-references: EMBL:X53462; NID:958978; PIDN:CAA37553.1; PID:958882  
C:Superfamily: SBV probable coat protein

Query Match 10.0%; Score 99.5; DB 2; Length 216;  
Best Local Similarity 25.0%; Pred. No. 0.083; Indels 29; Gaps 7;  
Matches 50; Conservative 27; Mismatches 94

QY 5 SDNLSNLSNVIT-DASSLNGVCKLLSAEVEKMLVQKAPNEGIEVFGLLYALAAARTSP 64  
DB 34 SEVNPKNLRKETDELGLVIRERFSSELV-----ITDEDFVHKLAFALIRAAINITSV 86  
QY 65 KVQASDVIFNSFGERNVVTGGDLKVLGCGAPLRTFTNKLRTFGRTTTEYVDFCI 124  
DB 87 KVNIVGA---YEYITGGKFLVKDAMVFFLIKECMKFKNPNVPTFCATPEDAY----I 139  
QY 125 AYKHKLQPL---NAAALGIPAEDSYLAADF--GTC-----PKLSE:QQSRKMFASKYA 174  
DB 140 VIARSAPKFLNATIKGIPSGYEFGLGADFLNTSVCLNDHEKAIVLQASRAIDRAVS 199  
QY 175 LKTEGGVNTVPSNLRQLQR 194

200 SSVGKIV-----SLFDLGR 214  
RESULT 8  
A33988  
adenylate cyclase (EC 4.6.1.1) - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 23-Mar-1990 #sequence\_revision 23-Mar-1990 #text\_change 31-Mar-2000  
C:Accession: A33988; A33539; T39809  
R:Young, D.; Riggs, M.; Field, J.; Vojtek, A.; Broek, D.; Wigler, M.  
Proc. Natl. Acad. Sci. U.S.A. 86, 7989-7993, 1989  
A:Title: The adenylate cyclase gene from Schizosaccharomyces pombe.  
A:Reference number: A33988; MUID:90046723; PMID:2682634  
A:Accession: A33988  
A:Molecule type: DNA  
A:Residues: 1-1692 <YOO>  
A:Cross-references: GB:M26699; NID:gl73338; PIDN:AAA35284.1; PID:gl73339  
R:Yamawaki-Kataoka, Y.; Tamacki, T.; Choe, H.R.; Tanaka, H.; Kataoka, T.  
Proc. Natl. Acad. Sci. U.S.A. 86, 5693-5697, 1989  
A:Title: Adenylate cyclases in yeast: a comparison of the genes from Schizosaccharomyces  
A:Reference number: A33539; MUID:89345533; PMID:2668944  
A:Accession: A33539  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1692 <YAM>  
A:Cross-references: GB:M24942; NID:gl73378; PIDN:AAA35301.1; PID:gl73379  
A:Note: the authors translated the codon TGC for residue 626 as Ser, and GCC for residu  
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lucas, M.; Gailardin, C.  
submitted to the EMBL data library, June 1998  
A:Reference number: 221981  
A:Accession: T39809  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-1692 <LYN>  
A:Cross-references: EMBL:AL023859; PIDN:CAA19571.1; GSPDB:GN00067; SPDS:SPRC19C7.03  
A:Experimental source: strain 972n-; cosmid c19C7  
C:Genetics:  
A:Gene: SPBC19C7.03  
A:Map position: 2  
C:Superfamily: yeast adenylate cyclase; leucine-rich alpha-2-glycoprotein repeat homolo  
C:KeyWords: phosphorus-oxygen lyase  
F:1328-1413/Domain: yeast adenylate cyclase catalytic domain homology <YACC>

Query Match 9.3%; Score 92; DB 2; Length 1692;  
Best Local Similarity 24.7%; Pred. No. 5.9;  
Matches 42; Conservative 26; Mismatches 68; Indels 34; Gaps 7;

QY 27 LLSAEVEKMLVQKAPNEG:RWVFGLLYALAAARTTSFKVQRADSDVIFNSFGERNVVV 86  
DB 339 LIQFNTERILL-----PHEQPCIIIFERLLSLFGCKVTS-----DEE:NESDNIYVARLVF 388  
QY 87 TEGDLKVLGCGAPLRTFTNKLRT-----FQRTFTTEYVDFCIAYKHKLQPLNAAALG:G 142  
DB 389 TMDI-----GADVLRKFEKKITANLDISRSNLEVIPVKIYPAHELT:SLNVSHNLSLD 443  
QY 143 AEDSYLAADF-GTCPKLSE:QQSRKMFASMLKTEGGVNTVPSNLRQL 192  
DB 444 -----LPLDFEVCVKLKRKIDISN-----LRSPRG---KPITAEURL 478

RESULT 9  
AD3246  
p-450 monooxygenase virH1 [imported] - Agrobacterium tumefaciens (strain C58, Dupont) P  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C:Accession: AD3246  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2322, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21605550; PMID:11743193

A:Accession: AD3246

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-419 <K>

A:Cross-references: GB:AE008690; PIDN:AAL46386.1; P.D:G:7744179; GSPDB:GN00169

A:Experimental source: strain C58 (Dipert)

C:Genetics:

A:Gene: virH

A:Genome: plasmid

C:Superfamily: Agrobacterium plasmid cytochrome P450 pinF1; cytochrome P450 homology

C:Keywords: heme; iron; metalloprotein

F:366/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 9.1%; Score 90; DB 2; Length 419;  
Best Local Similarity 25.1%; Pred. No. 1.5;  
Matches 52; Conservative 26; Mismatches 71; Indels 58; Gaps 11;

QY 20 LNVGDKLLSAEVEKMLVQKAPNEGIE--VVGILLLYALAARTTSPKVQADSDVIFSN 77  
DB 67 LIGTPRTQRIETELML-----NGVTRGAVDFDIRYSLFNSGEVHVKRSA---FAK 117  
QY 78 SFGERNVVVTEGLKXVLDGCA--LTRFTNKLRTFGRTFTAYVDFCIAYKHKLPQLNAA 136  
DB 119 TEAFR-----MDALRPEITKLTLEHLWDVPRVDD--PDFAEWYASKLPALTA 164  
QY 137 RELGIPAE-----SYLAADFLCTCPKLSELO-----OSRKM--- 168  
DB 165 SVLGPFPGDAPFTFLVNVSCSPSGWGEDDFPEASAVELODYRAVAVDRSRISD 224  
QY 169 -FASMY--ALKTEGGVWNTPVSNRLQL 192  
DB 225 DFLCYLKAVREGTL--SPIEIVQL 249

RESULT 10

F69673

C:phosphoenolpyruvate carboxykinase pckA - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000

C:Accession: F69673

R:Kunat, F.; Cgasawara, N.; Moszer, I.; Albertins, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A:Authors: Krogh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoesappe, S.; Hosono, S.; Hu

A,Map position: 2  
C,Superfamily: myrosinase-associated protein MyAP

Query Match 8.7%; Score 86; DB 2; Length 360;  
Best Local Similarity 26.5%; Pred. No. 2.9;  
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLVITDASSINGV--DKLLSAEVEKMLVQKAPNEG1---EVVFGLLLYALAAAR 60  
DB 137 SQPNMFKSVIARLKGIVGKXAXEILNNAFVVSNGPNCDFILNYEIPSRLEVPFISG 196  
QY 61 TTSPPKVRADSDVIFSNFSGERNVWVTEGLKKVLDGCAPLTRFNKLRTFGRTTEAYV 120  
DB 197 YQCFILKRLNFVRESYSLGVANLV--GGLPM--GCUPI--EMTAKFNIFRSCLEHNN 251  
QY 121 DFCIAYKHKLPOLNAAE--GIPAEYSYLAADFLGCPK;SELQOSRKMPASMYAK-TEG 179  
DB 252 KQSVLYNEKLQNLJPOIEASJFG-SKFLYADVYN--PMVEMIQNPSK-----YGFKE-TR 303  
QY 180 GVVNT 184  
DB 304 GCGGT 308

Query Match 8.7%; Score 86; DB 2; Length 447;  
Best Local Similarity 26.0%; Pred. No. 3.9;  
Matches 40; Conservative 15; Mismatches 41; Indels 58; Gaps 8;

QY 26 KLLSAEVEKMLVQKAPNEGIEVVVFGLLLYALAAARTTSPKVRADSDVIFSNFSGERNV- 84  
DB 5 KLESFIVEKMAERK-VFG-SISII-----KQGDVVYAKGGYERNE 44  
QY 85 -----VTEGLKKVLDGCAPLTRFN- KLRTFGRTT 116  
DB 45 ARLPSTPETTYGIGSITKSTFTALAIKMLVEGGLS--LDD--PVEKFN--KLRPFGPSVT 100  
QY 117 EAYVDFCIAYKHKLPOLNAAEEL--GIPAEOSYL 148  
DB 101 ---VHLLTHSSGIFSLGYAEAFIDGMVGGGNWL 131

RESULT 15  
T08639  
hypothetical protein b2627 - Escherichia coli (strain K-12)  
C,Species: Escherichia coli  
C,Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 01-Mar-2002  
C,Accession: T08639; E65041  
R,Plunkett, G.  
submitted to the EMBL Data Library, September 1995  
A,Reference number: Z16465  
A,Accession: T08639  
A,Status: preliminary; translated from GB/EMBL/DBJ  
A,Molecule type: DNA  
A,Residues: 1-729 <PLU>  
A,Cross-references: EMBL:U16840; NID:gl033110; PID:g1033122  
A,Experimental source: strain K12, substrain MGL655  
R,Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A,Title: The complete genome sequence of Escherichia coli K-12.  
A,Reference number: A64720; MUID:97426617; PMID:9278503  
A,Accession: E65041  
A,Status: preliminary; nucleic acid sequence not shown; translation not shown  
A,Molecule type: DNA  
A,Residues: 1-729 <BLAT>  
A,Cross-references: GB:AE000348; GB:U003096; NID:gl788975; PIDN:AAC75675.1; PID:g788980  
A,Experimental source: strain K-12, substrain MG1655  
C,Genetics:  
A,Map position: 57 min

Query Match 8.6%; Score 85.5; DB 2; Length 729;  
Best Local Similarity 22.4%; Pred. No. 8.1;  
Matches 41; Conservative 31; Mismatches 64; Indels 47; Gaps 8;

QY 2 ELMDSNLSNLVI--TDASSINGVDKLLSAEVEKMLVQKAPNEGIEVVFGLLLYALAA 59  
DB 520 EMDNDSNLSNLVI--TDASSINGVDKLLSAEVEKMLVQKAPNEGIEVVFGLLLYALAA 579  
QY 60 RTTSPKVRADSDVIFSNFSGERNV--VTEGLKKVLDGCAPLTRFNKLRTFGRT 114  
DB 580 ESYSEYKL---NQIDYARQWISGEKRTLSIALNNDUKL-----TN---TFGYT 623  
QY 115 FTEA-----YVDFCIAYK--HKLPOLNAAEELGIPAEOSYLAA 150

A,Map position: 2  
C,Superfamily: myrosinase-associated protein MyAP

Query Match 8.7%; Score 86; DB 2; Length 360;  
Best Local Similarity 26.5%; Pred. No. 2.9;  
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLVITDASSINGV--DKLLSAEVEKMLVQKAPNEG1---EVVFGLLLYALAAAR 60  
DB 137 SQPNMFKSVIARLKGIVGKXAXEILNNAFVVSNGPNCDFILNYEIPSRLEVPFISG 196  
QY 61 TTSPPKVRADSDVIFSNFSGERNVWVTEGLKKVLDGCAPLTRFNKLRTFGRTTEAYV 120  
DB 197 YQCFILKRLNFVRESYSLGVANLV--GGLPM--GCUPI--EMTAKFNIFRSCLEHNN 251  
QY 121 DFCIAYKHKLPOLNAAE--GIPAEYSYLAADFLGCPK;SELQOSRKMPASMYAK-TEG 179  
DB 252 KQSVLYNEKLQNLJPOIEASJFG-SKFLYADVYN--PMVEMIQNPSK-----YGFKE-TR 303  
QY 180 GVVNT 184  
DB 304 GCGGT 308

Query Match 8.7%; Score 86; DB 2; Length 445;  
Best Local Similarity 24.7%; Pred. No. 3.8;  
Matches 54; Conservative 41; Mismatches 80; Indels 44; Gaps 12;

QY 1 MELMSNLSNLVITDAS-----SLNGVDKLLSAEVEKMLVQKAPNEGIEV----- 48  
DB 32 IKILNDSNLSNLVITDNGKFSVIGNLGAENKLGEGVEMENLKIDNSNGGIAIPGFVD 91  
QY 49 -----VF-GLLLYALA--ARTSPKVRADSDVIFSNFSGERNVWVTEGLKKVLDGCA 99  
DB 92 GHSHPVFGSDRVEHFAPKLAGATYMEVQAGGIFPTTN---KTREASSQDKRDFEELA 148  
QY 100 PLTRFNKLRTFGRTTEAYVDFCIAYKHKLPOLN--AAAEELGIPAEOSYLAACFLG--T 155  
DB 149 -----KMLRSGLTTLEAKSGVGLNVDKXKMLRVLATENPNIPJE-----VSATTCGAHA 199  
QY 156 CPKLS-ELQOSR-----KMFASMYALKTEGVVNTPVSNL 189  
DB 200 VPKGSYEYQTRMICEELIPKTEDEKRNGLKN--VENI 236

RESULT 14  
G75201  
probable beta-lactamase (EC 3.5.2.6) PAB0087 - Pyrococcus abyssi (strain Ocsay)  
N,Alternate names: penicillin-binding protein homolog  
C,Species: Pyrococcus abyssi  
C,Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 05-Nov-1999  
C,Accession: G75201

624 LKVL\$MEDVVKH\$H\$V\$K\$G\$R\$S\$K\$V\$D\$Y\$T\$H\$V\$K\$L\$A\$F\$E\$S\$H\$U\$P\$G\$W\$A\$L\$E\$B\$I\$G\$P\$I\$Q\$T\$H\$R\$J\$V 683  
QY 15: DFL 153  
Db 684 DLL 686

Search completed: November 7, 2003, 14:50:22  
Job time : 23 secs

Result No.	Score	Query		ID	Description
		Match	Length		
1	125	12.6	223	COAT_STV36	Q0686 citrus t3ls
2	92	9.3	1692	CVAA_SCHPO	P14605 schizosach
3	89.5	9.0	397	AXZ_YARLI	Q9289 yarrowia li
4	87	8.8	527	PPCK_BACSU	P54418 bacillus su
5	85.5	8.6	729	YFKC_SCOCU	P52126 escherichia
6	84	8.5	537	PPCK_RHOFA	Q9284 rhodospirillum
7	83.5	8.4	483	TLDD_BUCAL	P57478 buchnera ap
8	83	8.4	395	ARGJ_LEPIN	Q8eyv8 l arginine
9	83	8.4	536	PPCK_AGR75	Q8ui94 agrobacteri
10	81.5	8.2	401	ACNA_THETN	Q8f9v4 thermoanaer
11	80.5	8.2	813	LOM2_BORBU	Q51558 borella bu
12	80	8.1	366	SPSY_MOUSE	P97355 mus musculu
13	79	8.0	308	OTCA_SYNY3	Q55497 synechocyst
14	79	8.0	422	CPXC_AGR7U	P24466 agrobacteri
15	79	8.0	527	PPCK_FUSNN	Q8tei2 fusobacteri
16	79	8.0	536	PPCK_RHIVE	P43085 rhizobium m
17	79	8.0	609	Y4PA_RHISN	P55610 rhizobium s
18	78.5	7.9	209	PYRE_LACLA	Q9Cgm8 lactococcus
19	78	7.9	477	DUD2_PSEPU	P31052 pseudomonas
20	76.5	7.7	500	H1X2_RHILC	Q988p7 rhizobium l
21	76.5	7.7	641	CAN6_RAT	Q88501 rattus norv
22	76	7.7	1016	IF3C_ARATH	Q9shil arabidopsis
23	75.5	7.6	538	PPCK_BRAIN	Q98bk7 bradyrhizob
24	75	7.6	277	ARGE_LEPBI	P13440 leptospira
25	75	7.6	536	PPCK_BRUSE	Q8ye41 bruceella me
26	75	7.6	526	PPCK_BRUSU	Q9fy05 bruceella su
27	75	7.6	883	CLPB_SVNP7	P53533 synechococc
28	74.5	7.5	222	Y175_METJA	P57639 methanococc
29	74.5	7.5	359	H158_SALT1	Q8f5j9 salmonella
30	74.5	7.5	434	PURB_WIGBR	Q84322 wiggleswort
31	74.5	7.5	581	AMV1_SCHPO	Q98420 schizosacch
32	74	7.5	366	SPSY_HUMAN	P52788 homo sapien
33	73.5	7.4	262	YY00_MYCTU	Q50725 mycobacteri

```

DB 89 VSSSLQSDDDATGITYTREGVEVLSOKLWTDVFNKSGIGNR----- 132
QY 99 APLTRFTNKLRTFGRTTAYVDFCIAYKHLPQLNAAAEIGIPAECSYLAADFGLGTCPK 158
DB 133 -----TNAIRVGRNTDALYLAFRCQRNRLSYGGRPLDAGIPAGYHYLCDFLTCAG 184
QY 159 LSELQSRKXKFSMAVXALTEGGVWVTPVSNLRQLQR 194
DB 185 LTDLECAVYIQAEOLK-KRGADVWVTVNVAQJGK 219

RESULT 2
CYAA SCHPO STANDARD; PRT; 1692 AA.
AC AC
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adenylylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl
DE cyclase).
GN CYRI OR SPBC19C7.03.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=90046723; PubMed=2682634;
RRT Young D., Riggs M., Field J., Vojtek A., Broek D., Wigler M.;
RRT "The adenylyl cyclase gene from Schizosaccharomyces pombe.";
RRL Proc. Natl. Acad. Sci. U.S.A. 86:7989-7993(1989).
RRL [2]
RP SEQUENCE FROM N.A.
MEDLINE=89345333; PubMed=2668944;
RRT Yamawaki-Kataoka Y., Tanacki T., Choe H.-R., Tanaka H., Kataoka T.;
RRT "Adenylylate cyclases in yeast: a comparison of the genes from
RRT Schizosaccharomyces pombe and Saccharomyces cerevisiae.";
RRL Proc. Natl. Acad. Sci. U.S.A. 96:5693-5697(1999).
RRL [3]
RP SEQUENCE FROM N.A.
STRAIN=972;
MEDLINE=21848401; PubMed=11859360;
RRT Wood V., Williams R., Rajandream M.A., Lyne Y., Lyne R., Stewart A.,
RRT Sgouros J., Peat N., Hayles C., Baker S., Basham D., Bowman S.,
RRT Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RRT Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RRT Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RRT Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jages K.,
RRT James K., Jones D., Jones M., Leather S., McDorald S., McLean J.,
RRT Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RRT Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RRT Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RRT Skerton J., Simmonds M., Squares K., Squares S., Stevens K.,
RRT Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RRT Woodward J., Volkhardt G., Aert R., Robben C., Grymoprez B.,
RRT Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RRT Gabel C., Fuchs M., Fritz C., Holzer E., Voestl D., Hilbert H.,
RRT Borzym K., Langer I., Beck A., Leirach K., Reinhardt R., Pohl T.M.,
RRT Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RRT Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaue K., Mottier S.,
RRT Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RRT Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RRT Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RRT Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.E.,
RRT Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RRT Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RRL "The genome sequence of Schizosaccharomyces pombe.";
RRL Nature 415:871-880(2002).
CC -!- FUNCTION: Plays essential roles in regulation of cellular
CC metabolism by catalyzing the synthesis of a second messenger,
CC cAMP.

```

```

CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -!- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -!- ENZYME REGULATION: In contrast to yeast cyclase, S.pombe cyclase
CC is not likely to be regulated by RAS proteins.
CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-3 family.
CC -!- SIMILARITY: Contains 21 leucine-rich (LRR) repeats.
CC -!- SIMILARITY: Contains 1 P2C-like domain.
CC -!- SIMILARITY: Contains 1 Ras-associating domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M26699; AAA35284.1; -
CC EMBL; M24942; AAA35301.1; -
CC EMBL; AL023859; CAA19571.1; -
CC PIR; A33988; A33988.
CC
CC GeneDB SPombe; SPBC19C7.03; -
CC InterPro; IPR001054; G_cyclase.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR003591; LRR_Typ.
CC InterPro; IPR001932; P2C-like.
CC InterPro; IPR000159; RA_Domain.
CC Pfam; PF00211; guanylate_cyc; 1.
CC Pfam; PF00560; LRR; 11.
CC Pfam; PF00481; P2C; 1.
CC SMART; SM00044; CYCC; 1.
CC SMART; SM00369; LRR_TYP; 1.
CC SMART; SM00332; P2C; 1.
CC SMART; SM00314; RA; 1.
CC PROSITE; PS0125; GUANYLATE_CYCLASES_2; 1.
CC PROSITE; PS0260; RA; 1.
CC Lyase; Repeat; Leucine-rich repeat; cAMP biosynthesis; Metal-binding;
CC Magnesium.
CC DOMAIN 292 380 RAS-ASSOCIATING.
CC REPEAT 385 403 LRR 1.
CC REPEAT 404 427 LRR 2.
CC REPEAT 429 451 LRR 3.
CC REPEAT 452 475 LRR 4.
CC REPEAT 476 498 LRR 5.
CC REPEAT 501 524 LRR 6.
CC REPEAT 525 547 LRR 7.
CC REPEAT 549 570 LRR 8.
CC REPEAT 572 594 LRR 9.
CC REPEAT 595 618 LRR 10.
CC REPEAT 619 641 LRR 11.
CC REPEAT 642 665 LRR 12.
CC REPEAT 666 689 LRR 13.
CC REPEAT 690 713 LRR 14.
CC REPEAT 714 737 LRR 15.
CC REPEAT 738 761 LRR 16.
CC REPEAT 762 785 LRR 17.
CC REPEAT 786 809 LRR 18.
CC REPEAT 810 833 LRR 19.
CC REPEAT 834 857 LRR 20.
CC REPEAT 858 881 LRR 21.
CC REPEAT 882 905 LRR 22.
CC REPEAT 906 929 LRR 23.
CC REPEAT 930 953 LRR 24.
CC REPEAT 954 977 LRR 25.
CC REPEAT 978 1001 LRR 26.
CC REPEAT 1002 1025 LRR 27.
CC REPEAT 1026 1049 LRR 28.
CC REPEAT 1050 1073 LRR 29.
CC REPEAT 1074 1097 LRR 30.
CC REPEAT 1098 1121 LRR 31.
CC REPEAT 1122 1145 LRR 32.
CC REPEAT 1146 1169 LRR 33.
CC REPEAT 1170 1193 LRR 34.
CC REPEAT 1194 1217 LRR 35.
CC REPEAT 1218 1241 LRR 36.
CC REPEAT 1242 1265 LRR 37.
CC REPEAT 1266 1289 LRR 38.
CC REPEAT 1290 1313 LRR 39.
CC REPEAT 1314 1337 LRR 40.
CC REPEAT 1338 1361 LRR 41.
CC REPEAT 1362 1385 LRR 42.
CC REPEAT 1386 1409 LRR 43.
CC REPEAT 1410 1433 LRR 44.
CC REPEAT 1434 1457 LRR 45.
CC REPEAT 1458 1481 LRR 46.
CC REPEAT 1482 1505 LRR 47.
CC REPEAT 1506 1529 LRR 48.
CC REPEAT 1530 1553 LRR 49.
CC REPEAT 1554 1577 LRR 50.
CC REPEAT 1578 1601 LRR 51.
CC REPEAT 1602 1625 LRR 52.
CC REPEAT 1626 1649 LRR 53.
CC REPEAT 1650 1673 LRR 54.
CC REPEAT 1674 1697 LRR 55.
CC REPEAT 1698 1721 LRR 56.
CC REPEAT 1722 1745 LRR 57.
CC REPEAT 1746 1769 LRR 58.
CC REPEAT 1770 1793 LRR 59.
CC REPEAT 1794 1817 LRR 60.
CC REPEAT 1818 1841 LRR 61.
CC REPEAT 1842 1865 LRR 62.
CC REPEAT 1866 1889 LRR 63.
CC REPEAT 1890 1913 LRR 64.
CC REPEAT 1914 1937 LRR 65.
CC REPEAT 1938 1961 LRR 66.
CC REPEAT 1962 1985 LRR 67.
CC REPEAT 1986 2009 LRR 68.
CC REPEAT 2010 2033 LRR 69.
CC REPEAT 2034 2057 LRR 70.
CC REPEAT 2058 2081 LRR 71.
CC REPEAT 2082 2105 LRR 72.
CC REPEAT 2106 2129 LRR 73.
CC REPEAT 2130 2153 LRR 74.
CC REPEAT 2154 2177 LRR 75.
CC REPEAT 2178 2201 LRR 76.
CC REPEAT 2202 2225 LRR 77.
CC REPEAT 2226 2249 LRR 78.
CC REPEAT 2250 2273 LRR 79.
CC REPEAT 2274 2297 LRR 80.
CC REPEAT 2298 2321 LRR 81.
CC REPEAT 2322 2345 LRR 82.
CC REPEAT 2346 2369 LRR 83.
CC REPEAT 2370 2393 LRR 84.
CC REPEAT 2394 2417 LRR 85.
CC REPEAT 2418 2441 LRR 86.
CC REPEAT 2442 2465 LRR 87.
CC REPEAT 2466 2489 LRR 88.
CC REPEAT 2490 2513 LRR 89.
CC REPEAT 2514 2537 LRR 90.
CC REPEAT 2538 2561 LRR 91.
CC REPEAT 2562 2585 LRR 92.
CC REPEAT 2586 2609 LRR 93.
CC REPEAT 2610 2633 LRR 94.
CC REPEAT 2634 2657 LRR 95.
CC REPEAT 2658 2681 LRR 96.
CC REPEAT 2682 2705 LRR 97.
CC REPEAT 2706 2729 LRR 98.
CC REPEAT 2730 2753 LRR 99.
CC REPEAT 2754 2777 LRR 100.
CC REPEAT 2778 2801 LRR 101.
CC REPEAT 2802 2825 LRR 102.
CC REPEAT 2826 2849 LRR 103.
CC REPEAT 2850 2873 LRR 104.
CC REPEAT 2874 2897 LRR 105.
CC REPEAT 2898 2921 LRR 106.
CC REPEAT 2922 2945 LRR 107.
CC REPEAT 2946 2969 LRR 108.
CC REPEAT 2970 2993 LRR 109.
CC REPEAT 2994 3017 LRR 110.
CC REPEAT 3018 3041 LRR 111.
CC REPEAT 3042 3065 LRR 112.
CC REPEAT 3066 3089 LRR 113.
CC REPEAT 3090 3113 LRR 114.
CC REPEAT 3114 3137 LRR 115.
CC REPEAT 3138 3161 LRR 116.
CC REPEAT 3162 3185 LRR 117.
CC REPEAT 3186 3209 LRR 118.
CC REPEAT 3210 3233 LRR 119.
CC REPEAT 3234 3257 LRR 120.
CC REPEAT 3258 3281 LRR 121.
CC REPEAT 3282 3305 LRR 122.
CC REPEAT 3306 3329 LRR 123.
CC REPEAT 3330 3353 LRR 124.
CC REPEAT 3354 3377 LRR 125.
CC REPEAT 3378 3401 LRR 126.
CC REPEAT 3402 3425 LRR 127.
CC REPEAT 3426 3449 LRR 128.
CC REPEAT 3450 3473 LRR 129.
CC REPEAT 3474 3497 LRR 130.
CC REPEAT 3498 3521 LRR 131.
CC REPEAT 3522 3545 LRR 132.
CC REPEAT 3546 3569 LRR 133.
CC REPEAT 3570 3593 LRR 134.
CC REPEAT 3594 3617 LRR 135.
CC REPEAT 3618 3641 LRR 136.
CC REPEAT 3642 3665 LRR 137.
CC REPEAT 3666 3689 LRR 138.
CC REPEAT 3690 3713 LRR 139.
CC REPEAT 3714 3737 LRR 140.
CC REPEAT 3738 3761 LRR 141.
CC REPEAT 3762 3785 LRR 142.
CC REPEAT 3786 3809 LRR 143.
CC REPEAT 3810 3833 LRR 144.
CC REPEAT 3834 3857 LRR 145.
CC REPEAT 3858 3881 LRR 146.
CC REPEAT 3882 3905 LRR 147.
CC REPEAT 3906 3929 LRR 148.
CC REPEAT 3930 3953 LRR 149.
CC REPEAT 3954 3977 LRR 150.
CC REPEAT 3978 4001 LRR 151.
CC REPEAT 4002 4025 LRR 152.
CC REPEAT 4026 4049 LRR 153.
CC REPEAT 4050 4073 LRR 154.
CC REPEAT 4074 4097 LRR 155.
CC REPEAT 4098 4121 LRR 156.
CC REPEAT 4122 4145 LRR 157.
CC REPEAT 4146 4169 LRR 158.
CC REPEAT 4170 4193 LRR 159.
CC REPEAT 4194 4217 LRR 160.
CC REPEAT 4218 4241 LRR 161.
CC REPEAT 4242 4265 LRR 162.
CC REPEAT 4266 4289 LRR 163.
CC REPEAT 4290 4313 LRR 164.
CC REPEAT 4314 4337 LRR 165.
CC REPEAT 4338 4361 LRR 166.
CC REPEAT 4362 4385 LRR 167.
CC REPEAT 4386 4409 LRR 168.
CC REPEAT 4410 4433 LRR 169.
CC REPEAT 4434 4457 LRR 170.
CC REPEAT 4458 4481 LRR 171.
CC REPEAT 4482 4505 LRR 172.
CC REPEAT 4506 4529 LRR 173.
CC REPEAT 4530 4553 LRR 174.
CC REPEAT 4554 4577 LRR 175.
CC REPEAT 4578 4601 LRR 176.
CC REPEAT 4602 4625 LRR 177.
CC REPEAT 4626 4649 LRR 178.
CC REPEAT 4650 4673 LRR 179.
CC REPEAT 4674 4697 LRR 180.
CC REPEAT 4698 4721 LRR 181.
CC REPEAT 4722 4745 LRR 182.
CC REPEAT 4746 4769 LRR 183.
CC REPEAT 4770 4793 LRR 184.
CC REPEAT 4794 4817 LRR 185.
CC REPEAT 4818 4841 LRR 186.
CC REPEAT 4842 4865 LRR 187.
CC REPEAT 4866 4889 LRR 188.
CC REPEAT 4890 4913 LRR 189.
CC REPEAT 4914 4937 LRR 190.
CC REPEAT 4938 4961 LRR 191.
CC REPEAT 4962 4985 LRR 192.
CC REPEAT 4986 5009 LRR 193.
CC REPEAT 5010 5033 LRR 194.
CC REPEAT 5034 5057 LRR 195.
CC REPEAT 5058 5081 LRR 196.
CC REPEAT 5082 5105 LRR 197.
CC REPEAT 5106 5129 LRR 198.
CC REPEAT 5130 5153 LRR 199.
CC REPEAT 5154 5177 LRR 200.
CC REPEAT 5178 5201 LRR 201.
CC REPEAT 5202 5225 LRR 202.
CC REPEAT 5226 5249 LRR 203.
CC REPEAT 5250 5273 LRR 204.
CC REPEAT 5274 5297 LRR 205.
CC REPEAT 5298 5321 LRR 206.
CC REPEAT 5322 5345 LRR 207.
CC REPEAT 5346 5369 LRR 208.
CC REPEAT 5370 5393 LRR 209.
CC REPEAT 5394 5417 LRR 210.
CC REPEAT 5418 5441 LRR 211.
CC REPEAT 5442 5465 LRR 212.
CC REPEAT 5466 5489 LRR 213.
CC REPEAT 5490 5513 LRR 214.
CC REPEAT 5514 5537 LRR 215.
CC REPEAT 5538 5561 LRR 216.
CC REPEAT 5562 5585 LRR 217.
CC REPEAT 5586 5609 LRR 218.
CC REPEAT 5610 5633 LRR 219.
CC REPEAT 5634 5657 LRR 220.
CC REPEAT 5658 5681 LRR 221.
CC REPEAT 5682 5705 LRR 222.
CC REPEAT 5706 5729 LRR 223.
CC REPEAT 5730 5753 LRR 224.
CC REPEAT 5754 5777 LRR 225.
CC REPEAT 5778 5801 LRR 226.
CC REPEAT 5802 5825 LRR 227.
CC REPEAT 5826 5849 LRR 228.
CC REPEAT 5850 5873 LRR 229.
CC REPEAT 5874 5897 LRR 230.
CC REPEAT 5898 5921 LRR 231.
CC REPEAT 5922 5945 LRR 232.
CC REPEAT 5946 5969 LRR 233.
CC REPEAT 5970 5993 LRR 234.
CC REPEAT 5994 6017 LRR 235.
CC REPEAT 6018 6041 LRR 236.
CC REPEAT 6042 6065 LRR 237.
CC REPEAT 6066 6089 LRR 238.
CC REPEAT 6090 6113 LRR 239.
CC REPEAT 6114 6137 LRR 240.
CC REPEAT 6138 6161 LRR 241.
CC REPEAT 6162 6185 LRR 242.
CC REPEAT 6186 6209 LRR 243.
CC REPEAT 6210 6233 LRR 244.
CC REPEAT 6234 6257 LRR 245.
CC REPEAT 6258 6281 LRR 246.
CC REPEAT 6282 6305 LRR 247.
CC REPEAT 6306 6329 LRR 248.
CC REPEAT 6330 6353 LRR 249.
CC REPEAT 6354 6377 LRR 250.
CC REPEAT 6378 6401 LRR 251.
CC REPEAT 6402 6425 LRR 252.
CC REPEAT 6426 6449 LRR 253.
CC REPEAT 6450 6473 LRR 254.
CC REPEAT 6474 6497 LRR 255.
CC REPEAT 6498 6521 LRR 256.
CC REPEAT 6522 6545 LRR 257.
CC REPEAT 6546 6569 LRR 258.
CC REPEAT 6570 6593 LRR 259.
CC REPEAT 6594 6617 LRR 260.
CC REPEAT 6618 6641 LRR 261.
CC REPEAT 6642 6665 LRR 262.
CC REPEAT 6666 6689 LRR 263.
CC REPEAT 6690 6713 LRR 264.
CC REPEAT 6714 6737 LRR 265.
CC REPEAT 6738 6761 LRR 266.
CC REPEAT 6762 6785 LRR 267.
CC REPEAT 6786 6809 LRR 268.
CC REPEAT 6810 6833 LRR 269.
CC REPEAT 6834 6857 LRR 270.
CC REPEAT 6858 6881 LRR 271.
CC REPEAT 6882 6905 LRR 272.
CC REPEAT 6906 6929 LRR 273.
CC REPEAT 6930 6953 LRR 274.
CC REPEAT 6954 6977 LRR 275.
CC REPEAT 6978 7001 LRR 276.
CC REPEAT 7002 7025 LRR 277.
CC REPEAT 7026 7049 LRR 278.
CC REPEAT 7050 7073 LRR 279.
CC REPEAT 7074 7097 LRR 280.
CC REPEAT 7098 7121 LRR 281.
CC REPEAT 7122 7145 LRR 282.
CC REPEAT 7146 7169 LRR 283.
CC REPEAT 7170 7193 LRR 284.
CC REPEAT 7194 7217 LRR 285.
CC REPEAT 7218 7241 LRR 286.
CC REPEAT 7242 7265 LRR 287.
CC REPEAT 7266 7289 LRR 288.
CC REPEAT 7290 7313 LRR 289.
CC REPEAT 7314 7337 LRR 290.
CC REPEAT 7338 7361 LRR 291.
CC REPEAT 7362 7385 LRR 292.
CC REPEAT 7386 7409 LRR 293.
CC REPEAT 7410 7433 LRR 294.
CC REPEAT 7434 7457 LRR 295.
CC REPEAT 7458 7481 LRR 296.
CC REPEAT 7482 7505 LRR 297.
CC REPEAT 7506 7529 LRR 298.
CC REPEAT 7530 7553 LRR 299.
CC REPEAT 7554 7577 LRR 300.
CC REPEAT 7578 7601 LRR 301.
CC REPEAT 7602 7625 LRR 302.
CC REPEAT 7626 7649 LRR 303.
CC REPEAT 7650 7673 LRR 304.
CC REPEAT 7674 7697 LRR 305.
CC REPEAT 7698 7721 LRR 306.
CC REPEAT 7722 7745 LRR 307.
CC REPEAT 7746 7769 LRR 308.
CC REPEAT 7770 7793 LRR 309.
CC REPEAT 7794 7817 LRR 310.
CC REPEAT 7818 7841 LRR 311.
CC REPEAT 7842 7865 LRR 312.
CC REPEAT 7866 7889 LRR 313.
CC REPEAT 7890 7913 LRR 314.
CC REPEAT 7914 7937 LRR 315.
CC REPEAT 7938 7961 LRR 316.
CC REPEAT 7962 7985 LRR 317.
CC REPEAT 7986 8009 LRR 318.
CC REPEAT 8010 8033 LRR 319.
CC REPEAT 8034 8057 LRR 320.
CC REPEAT 8058 8081 LRR 321.
CC REPEAT 8082 8105 LRR 322.
CC REPEAT 8106 8129 LRR 323.
CC REPEAT 8130 8153 LRR 324.
CC REPEAT 8154 8177 LRR 325.
CC REPEAT 8178 8201 LRR 326.
CC REPEAT 8202 8225 LRR 327.
CC REPEAT 8226 8249 LRR 328.
CC REPEAT 8250 8273 LRR 329.
CC REPEAT 8274 8297 LRR 330.
CC REPEAT 8298 8321 LRR 331.
CC REPEAT 8322 8345 LRR 332.
CC REPEAT 8346 8369 LRR 333.
CC REPEAT 8370 8393 LRR 334.
CC REPEAT 8394 8417 LRR 335.
CC REPEAT 8418 8441 LRR 336.
CC REPEAT 8442 8465 LRR 337.
CC REPEAT 8466 8489 LRR 338.
CC REPEAT 8490 8513 LRR 339.
CC REPEAT 8514 8537 LRR 340.
CC REPEAT 8538 8561 LRR 341.
CC REPEAT 8562 8585 LRR 342.
CC REPEAT 8586 8609 LRR 343.
CC REPEAT 8610 8633 LRR 344.
CC REPEAT 8634 8657 LRR 345.
CC REPEAT 8658 8681 LRR 346.
CC REPEAT 8682 8705 LRR 347.
CC REPEAT 8706 8729 LRR 348.
CC REPEAT 8730 8753 LRR 349.
CC REPEAT 8754 8777 LRR 350.
CC REPEAT 8778 8801 LRR 351.
CC REPEAT 8802 8825 LRR 352.
CC REPEAT 8826 8849 LRR 353.
CC REPEAT 8850 8873 LRR 354.
CC REPEAT 8874 8897 LRR 355.
CC REPEAT 8898 8921 LRR 356.
CC REPEAT 8922 8945 LRR 357.
CC REPEAT 8946 8969 LRR 358.
CC REPEAT 8970 8993 LRR 359.
CC REPEAT 8994 9017 LRR 360.
CC REPEAT 9018 9041 LRR 361.
CC REPEAT 9042 9065 LRR 362.
CC REPEAT 9066 9089 LRR 363.
CC REPEAT 9090 9113 LRR 364.
CC REPEAT 9114 9137 LRR 365.
CC REPEAT 9138 9161 LRR 366.
CC REPEAT 9162 9185 LRR 367.
CC REPEAT 9186 9209 LRR 368.
CC REPEAT 9210 9233 LRR 369.
CC REPEAT 9234 9257 LRR 370.
CC REPEAT 9258 9281 LRR 371.
CC REPEAT 9282 9305 LRR 372.
CC REPEAT 9306 9329 LRR 373.
CC REPEAT 9330 9353 LRR 374.
CC REPEAT 9354 9377 LRR 375.
CC REPEAT 9378 9401 LRR 376.
CC REPEAT 9402 9425 LRR 377.
CC REPEAT 9426 9449 LRR 378.
CC REPEAT 9450 9473 LRR 379.
CC REPEAT 9474 9497 LRR 380.
CC REPEAT 9498 9521 LRR 381.
CC REPEAT 9522 9545 LRR 382.
CC REPEAT 9546 9569 LRR 383.
CC REPEAT 9570 9593 LRR 384.
CC REPEAT 9594 9617 LRR 385.
CC REPEAT 9618 9641 LRR 386.
CC REPEAT 9642 9665 LRR 387.
CC REPEAT 9666 9689 LRR 388.
CC REPEAT 9690 9713 LRR 389.
CC REPEAT 9714 9737 LRR 390.
CC REPEAT 9738 9761 LRR 391.
CC REPEAT 9762 9785 LRR 392.
CC REPEAT 9786 9809 LRR 393.
CC REPEAT 9810 9833 LRR 394.
CC REPEAT 9834 9857 LRR 395.
CC REPEAT 9858 9881 LRR 396.
CC REPEAT 9882 9905 LRR 397.
CC REPEAT 9906 9929 LRR 398.
CC REPEAT 9930 9953 LRR 399.
CC REPEAT 9954 9977 LRR 400.
CC REPEAT 9978 10001 LRR 401.
CC REPEAT 10002 10025 LRR 402.
CC REPEAT 10026 10049 LRR 403.
CC REPEAT 10050 10073 LRR 404.
CC REPEAT 10074 10097 LRR 405.
CC REPEAT 10098 10121 LRR 406.
CC REPEAT 10122 10145 LRR 407.
CC REPEAT 10146 10169 LRR 408.
CC REPEAT 10170 10193 LRR 409.
CC REPEAT 10194 10217 LRR 410.
CC REPEAT 10218 10241 LRR 411.
CC REPEAT 10242 10265 LRR 412.
CC REPEAT 10266 10289 LRR 413.
CC REPEAT 10290 10313 LRR 414.
CC REPEAT 10314 10337 LRR 415.
CC REPEAT 10338 10361 LRR 416.
CC REPEAT 10362 10385 LRR 417.
CC REPEAT 10386 10409 LRR 418.
CC REPEAT 10410 10433 LRR 419.
CC REPEAT 10434 10457 LRR 420.
CC REPEAT 10458 10481 LRR 421.
CC REPEAT 10482 10505 LRR 422.
CC REPEAT 10506 10529 LRR 423.
CC REPEAT 10530 10553 LRR 424.
CC REPEAT 10554 10577 LRR 425.
CC REPEAT 10578 10601 LRR 426.
CC REPEAT 10602 10625 LRR 427.
CC REPEAT 10626 10649 LRR 428.
CC REPEAT 10650 10673 LRR 429.
CC REPEAT 10674 10697 LRR 430.
CC REPEAT 10698 10721 LRR 431.
CC REPEAT 10722 10745 LRR 432.
CC REPEAT 10746 10769 LRR 433.
CC REPEAT 10770 10793 LRR 434.
CC REPEAT 10794 10817 LRR 435.
CC REPEAT 10818 10841 LRR 436.
CC REPEAT 10842 10865 LRR 437.
CC REPEAT 10866 10889 LRR 438.
CC REPEAT 10890 10913 LRR 439.
CC REPEAT 10914 10937 LRR 440.
CC REPEAT 10938 10961 LRR 441.
CC REPEAT 10962 10985 LRR 442.
CC REPEAT 10986 11009 LRR 443.
CC REPEAT 11010 11033 LRR 444.
CC REPEAT 11034 11057 LRR 445.
CC REPEAT 11058 11081 LRR 446.
CC REPEAT 11082 11105 LRR 447.
CC REPEAT 11106 11129 LRR 448.
CC REPEAT 11130 11153 LRR 449.
CC REPEAT 11154 11177 LRR 450.
CC REPEAT 11178 11201 LRR 451.
CC REPEAT 11202 11225 LRR 452.
CC REPEAT 11226 11249 LRR 453.
CC REPEAT 11250 11273 LRR 454.
CC REPEAT 11274 11297 LRR 455.
CC REPEAT 11298 11321 LRR 456.
CC REPEAT 11322 11345 LRR 457.
CC REPEAT 11346 11369 LRR 458.
CC REPEAT 11370 11393 LRR 459.
CC REPEAT 11394 11417 LRR 460.
CC REPEAT 11418 11441 LRR 461.
CC REPEAT 11442 11465 LRR 462.
CC REPEAT 11466 11489 LRR 463.
CC REPEAT 11490 11513 LRR 464.
CC REPEAT 11514 11537 LRR 465.
CC REPEAT 11538 11561 LRR 466.
CC REPEAT 11562 11585 LRR 467.
CC REPEAT 11586 11609 LRR 468.
CC REPEAT 11610 11633 LRR 469.
CC REPEAT 11634 11657 LRR 470.
CC REPEAT 11658 11681 LRR 471.
CC REPEAT 11682 11705 LRR 472.
CC REPEAT 11706 11729 LRR 473.
CC REPEAT 11730 11753 LRR 474.
CC REPEAT 11754 11777 LRR 475.
CC REPEAT 11778 11801 LRR 476.
CC REPEAT 11802 11825 LRR 477.
CC REPEAT 11826 11849 LRR 478.
CC REPEAT 11850 11873 LRR 479.
CC REPEAT 11874 11897 LRR 
```

```

87  TEGDLKKVLDGCAPLTRFNKLT-----PGRFTTEAYVDFCIAYGKLPQLNAAAEICIP 142
389  TMDDI-----GADVURKSEKITANTLDIRSNLEVIPVKIYPYAHLEISUNVSHNSLD 443
143  AEDSYLAADFGLCTCPKSELQSRKXFASMYALKTEGGVWNTPTVSNRLQL 192
444  -----LPLDMERCVKLKELDISNN-----LRSPRG---KPIAJRQL 478

RESULT 3
AXP_YARLI
ID AXP_YARLI STANDARD; PRT: 397 AA.
AC Q92389;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acid extracellular protease precursor (EC 3.4.23.-).
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=148;
RC MEDLINE=97039861; PubMed=8885467;
RA Young T.W., Wadeson A., Glover D.J., Quincey R.V., Batlin M.J.,
RA Kamei E.A.;
RA "The extracellular acid protease gene of Yarrowia lipolytica:
RA sequence and pH-regulated transcription.";
RA Microbiology 142:2913-2921(1996);
RA -- SUBCELLULAR LOCATION: Secreted.
RA -- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
RA
RA This SWISS-PROT entry is copyrighted. It is produced through a collaboration:
RA between the Swiss Institute of Bioinformatics and the EMBL outstation -
RA the European Bioinformatics Institute. There are no restrictions on its
RA use by non-profit institutions as long as its content is in no way
RA modified and this statement is not removed. Usage by and for commercial
RA entities requires a license agreement (See http://www.isb-sib.ch/announce/
RA or send an email to license@isb-sib.ch).
RA
RA -----
RA EMBL; X97068; CAA65778.1; --
RA HSP: P07339; 1LYB.
RA MEROPS; A01.036; --
RA InterPro; IPR001969; AspproteaseA1.
RA InterPro; IPR001461; AspproteaseA1.
RA Pfam; PF00026; asp; 1.
RA PRINTS; PR00792; PEP5IN.
RA PROSITE; PS00141; ASP_PROTEASE; 1.
RA Hydrolase; Aspartyl protease; Zymogen; Signal.
RA SIGNAL 1 17
RA PROPEP 19 ?
RA CHAIN ? 397
RA ACT_SITE 77 77
RA BY SIMILARITY.
RA ACT_SITE 264 264
RA BY SIMILARITY.
RA DISULFID 93 100
RA BY SIMILARITY.
RA DISULFID 302 343
RA BY SIMILARITY.
RA CARBOHYD 88 88
RA N-LINKED (GLCNAC. .) (POTENTIAL).
RA CARBOHYD 310 310
RA N-LINKED (GLCNAC. .) (POTENTIAL).
RA CARBOHYD 314 314
RA N-LINKED (GLCNAC. .) (POTENTIAL).
RA SEQUENCE 397 AA; 4208; MW; BD678814B4B8984F CRC64;
RA
RA Query Match 9.0%; Score 89.5; DB 1; Length 397;
RA Best Local Similarity 21.9%; Pred. No. 0.91;
RA Matches 46; Conservative 37; Mismatches 86; Indels 41; Gaps 8;
RA
RA 15 TDASSLGVYKXLLSAEVEKMLVQXG-APNEG:EVVFGLLVALAARTSPKVGAD--S 71
RA 136 TDNAVIGGA-----SAPFKFGVNGDLSGGFSVMVFGVGNDSASTS:SAQLQKSGEIS 190
RA 72 DVIFNSFGERNVVVTEGSLKKVLGDCAPLRTFTNKLRFTFGRTTTE-AVYDFC:AYKHKL 130

```

```

191 RNLYGMSFSDANLACTGNDNSEITFGAINTGRTIGSLKTIPTVATOGGYOHFVSASGKF 250
131 PQE-----NAAAE:GIPAEYSYLAAD-----PLG-----TCPKLSLQ 163
251 GDVDFDFNDLVLDGSGTTTTLKSDYVNAFLGGLEDIDLTLSYSGWHGPGCSENSKIN 310
164 -----OSRKMFAWMYALKTEGGVWNTPTVSN 188
311 PTYNFSKGEITVTGHDLA:PGNAVNSNVDS 340

RESULT 4
PECK_BACSU
ID PECK_BACSU STANDARD; PRT: 527 AA.
AC P54418; C34304;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49; [PEP
DE carboxykinase] (Phosphoenolpyruvate carboxylase) (PEPCK).
OX PKA OR PPC.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=423;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=168;
RC MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RA "Sequencing and functional annotation of the Bacillus subtilis genes
RA in the 203 kb rnb-dnaB region.";
RA Microbiology 143:3431-3441(1997);
RA [2]
RA SEQUENCE FROM N.A.
RA STRAIN=168;
RA MEDLINE=9804033; PubMed=9384377;
RA Kunst F., Ogasawara N., Woszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Beasieres P., Bolotin A., Borchert S.,
RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codari J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Huilo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mael C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nockack M.,
RA Noone L., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Rapoport S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RA "The complete genome sequence of the Gram-positive bacterium Bacillus
RA subtilis.";
RA Nature 390:249-256(1997).
RA [3]
RA SEQUENCE OF 1-165 FROM N.A.
RA STRAIN=168 / PY79;
RA MEDLINE=96345628; PubMed=875891;
RA Yocum R., Perkins J.B., Howitt C.L., Pero J.;

```



"Cloning and characterization of the metE gene encoding S-adenosylmethionine synthetase from *Bacillus subtilis*.";  
J. Bacteriol. 178:4604-4610(1996).  
-!- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate + CO(2).  
-!- PATHWAY: Rate-limiting gluconeogenic enzyme.  
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
-!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase (ATP) family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: AF008220; AAC00377.1; -  
EMBL: Z99119; CAB15034.1; -  
EMBL: U52812; AAB17065.1; -  
PIR: F69673; F69673.  
HSSP: P22239; 1AYL.  
Subtilist: BG11841; pckA.  
HAMAP: MF\_00453; -; 1.  
InterPro: IPR001272; PEPCK\_ATP.  
Pfam: PF01293; PEPCK\_ATP; 1.  
ProDom: PD004723; PEPCK\_ATP; 1.  
TIGRFAMs: TIGR00224; pckA; 1.  
PROSITE: PS00532; PEPCK\_ATP; 1.  
Glucoseogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.  
NP\_BIND 233 240 ATP (BY SIMILARITY).  
CONFLICT 10 10 L > S (IN REF. 3).  
SEQUENCE 527 AA; 58300 MW; E55:EC80ZD15666 CRC64;

Query Match 8.8%; Score 87; DB 1; Length 527;  
Best Local Similarity 26.9%; Pred. No. 2.2;  
Matches 47; Conservative 24; Mismatches 70; Indels 34; Gaps 9;

QY 20 LNVGVKLLSAEVEKLV---OKGAPNGIEVVFGLLYALAAARTSPKVRASDVIFS 76  
DB 234 LSGTGKTTTASADPKLIGDDEHFGASDTGVFNIEG-GCYAKCIHLSBK----EPQIFNA 288  
2Y 77 NSFGE--RNVVVTGDLKKVLDGCAPLTFRTNKLRTFTTEAVVDFCIAYK-HKLPQL 133  
DB 289 TRFGVLENVVDDE-----TREAYDVSFTENTRA-----AYPIHMNNI 330  
QY 134 NAAELGIPAEPSYLAADFLGTCPLKLSLQSSRKX--FASKYALK---TEGGVNV 183  
DB 331 VTPSMAGHPSAIVPLTADAFGLVPEPSIKLTKEQVMYHFLSGVTSKLAGTERGVTS 385

RESULT 5  
YFJK\_ECOLI  
ID YFJK\_ECOLI STANDARD; PRT; 729 AA.  
AC P52126;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein: yfjk.  
GN YFJK OR B2627.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RC  
RP  
RT  
STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
Mau B., Shao Y.;

"The complete genome sequence of *Escherichia coli* K-12.";  
Science 277:1233-1238(1997).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U36840; AAA79736.1; -  
EMBL: AE006348; AAC75675.1; -  
PIR: T08639; T08639.  
EcoGene: EGI3197; yfjk.  
InterPro: IPR001410; DEAD.  
DR InterPro: IPR001650; Helicase\_C.  
DR Pfam: P500271; helicase\_C; 1.  
DR SMART: SM00487; DEXDC; 1.  
DR SMART: SM00490; HELIC\_C; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 729 AA; 83061 MW; CB65BCD1B36FBFAF CRC64;

Query Match 8.6%; Score 85.5; DB 1; Length 729;  
Best Local Similarity 22.4%; Pred. No. 4.4;  
Matches 41; Conservative 31; Mismatches 64; Indels 47; Gaps 8;

QY 2 ELKSGSNLSNLSV---TDASSLNGVDKLLSAEVEKLVOKGAPNGIEVVFGLLYALAA 59  
DB 520 EMDNEPSSLVFHWTPATNFKLTKFAKIIARLVPHPTFSRNGVPEKPTDWMIAKLAGYLSA 579  
QY 60 RTTSPKVRASDVIFS---SFGERNV--VVTGDLKKVLDGCAPLTFRTNKLRTFTGRT 114  
DB 580 ESYSEYK---NQIDYARQWISGEKRTLSIANDLKI-----TV---TFGYT 623  
QY 115 FTEA-----YVDFCIAYK--HKLPOLNAAELGIPAEPSYLA 150  
DB 624 LPKVLNEDVYGHVHAKRGIRSKVDYTHVKLAFSEFHLPPGVNALEEIGIPQTLHRLV 683  
QY 151 DFL 153  
DB 684 DLL 686

RESULT 6  
PECK\_RHOPA  
ID PECK\_RHOPA STANDARD; PRT; 537 AA.  
AC Q9ZNH4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 26-FEB-2003 (Rel. 41, Last annotation update)  
DE Phosphoenolpyruvate carboxykinase (ATP) (EC 4.1.1.49) (PEP  
DE carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).  
GN pckA.  
OS Rhodospseudomonas palustris.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Rhodospseudomonas.  
OX NCBI\_TaxID=1076;  
RN [1]  
RP  
RT  
STRAIN=7;  
RX MEDLINE=99235744; PubMed=10217755;  
RA Imai M., Nakata K., Roh J.H., Zahn K., Yukawa H.;  
PT "Molecular and functional characterization of the Rhodospseudomonas  
RT palustris no. 7 phosphoenolpyruvate carboxykinase gene."  
RL J. Bacteriol. 181:2689-2696(1999).  
CC -!- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate  
CC + CO(2).  
CC -!- PATHWAY: Rate-limiting gluconeogenic enzyme.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS DURING LOG PHASE  
CC WITH 10-20 FOLD REDUCTION AT ONSET OF STATIONARY PHASE.  
CC -!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [ATP]

```

DR Pfam: PF01523; PmbA_TldD; 1.
KW Complete proteome.
SQ SEQUENCE 483 AA; 52977 MW; A05CE98518720EBA CRC64;

      Query Match      8.4%; Score 83.5; DB 1; Length 483;
      Best Local Similarity 20.5%; Pred. No. 4.4;
      Matches 32; Conservative 26; Mismatches 53; Indels 45; Gaps 5;

QY      65 KVRADSDVIFNSFGERNVWTEGDLKKVLD-----GCAPLTR 103
      ||| : |||::: |||:
Cb      305 KNQRGSLSDCGTPOCKNLIENGILKKYMQKLNARLMGVKSTGNRRRESYSLPMPR 364
      ||| : |||: |||: |||:
QY      104 FTKNKLRTGRTFAVVDVFCIAVKHKLPOLNAAELEGIPAEQDSYLAADFLGTCTPKLSELQ 163
      ||| : |||: |||: |||:
Cb      365 MTN-----TYMLSGKSLDIIKSVDYGI-----YAVNPSG---GQVDIT 40:

QY      164 QSRKMFASYALKTEGVVNTPVSNLRQLGR-REVW 198
      : |||: |||: |||: |||:
Db      402 SGKVFSTSEAVLKNKGKVTPIKNTLLIGSGLEVM 437

```

ARGC_ID	ARGC_LEPIN	STANDARD;	PRT;	385 AA.
AC	Q8EYV8;			
DE	15-SEP-2003 (Rel. 42, Created)			
DE	15-SEP-2003 (Rel. 42, Last sequence update)			
DE	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Arginine biosynthesis bifunctional protein argJ [Includes: Glutamate			
DE	N-acetyltransferase (EC 2.3.1.35) [Ornithine acetyltransferase;			
DE	(Ornithine transacetylase) (OATase); Amino-acid acetyltransferase			
DE	(EC 2.3.1.1) (N-acetylglutamate synthase) (AGS)] [Contains: Arginine			
DE	biosynthesis bifunctional protein argJ alpha chain; Arginine			
DE	biosynthesis bifunctional protein argJ beta chain].			
GK	ARGJ OR LA4105.			
OS	Leptospira interrogans.			
OC	Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.			
OX	NCBI TaxId=173;			

SEQUENCE FROM N.A.  
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
 RC MEDLINE=22598143; PubMed=12712204;  
 RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,  
 RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,  
 RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,  
 RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,  
 RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,  
 RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,  
 Xu J.-G., Zhao G.-P.;  
 RT "unique physiological and pathogenic features of *Leptospira*  
 RT interogans" revealed by whole-genome sequencing.";  
 RC Nature 422:888-893 (2003).  
 CC -|- FUNCTION: Catalyzes two activities which are involved in the  
 CC cyclic version of arginine biosynthesis: the synthesis of  
 CC acetylglutamate from glutamate and acetyl-CoA, and of ornithine by  
 CC transacylation between acetylornithine and glutamate (By  
 CC similarity).  
 CC -|- CATALYTIC ACTIVITY: N(2)-acetyl-L-ornithine + L-glutamate = L-  
 CC ornithine + N-acetyl-L-glutamate.  
 CC -|- CATALYTIC ACTIVITY: Acetyl-CoA + L-glutamate = CoA + N-acetyl-L-  
 CC glutamate.  
 CC -|- PATHWAY: Arginine biosynthesis; first step.  
 CC -|- PATHWAY: Arginine biosynthesis; fifth step.  
 CC -|- SUBUNIT: Heterotetramer of two alpha and two beta chains (By  
 CC similarity).  
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -|- MISCELLANEOUS: Some bacteria possess a monofunctional argJ, i.e.,  
 CC capable of catalyzing only the fifth step of the arginine  
 CC biosynthetic pathway.  
 CC -|- SIMILARITY: Belongs to the argC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation

C the European Bioinformatics Institute. There are no restrictions on its  
C use by non-profit institutions as long as its content is in no way  
C modified and this statement is not removed. Usage by and for commercial  
C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

R EMBL; AE011564; AAN51303.1; -  
R HAMAP; MF\_01106; -; 1.  
R InterPro; IPR002813; ArgJ.  
R Pfam; PF01960; ArgJ; 1.  
R ProDom; PD004193; ArgJ; 1.  
R TIGRFAMs; TIGR00120; ArgJ; 1.  
R Arginine biosynthesis; Multifunctional enzyme; Transferase;  
R Acyltransferase; Complete proteome.  
T CHAIN 1 178  
T ARGININE BIOSYNTHESIS BIFUNCTIONAL  
T PROTEIN ARGJ ALPHA CHAIN (BY SIMILARITY).  
T ARGININE BIOSYNTHESIS BIFUNCTIONAL  
T PROTEIN ARGJ BETA CHAIN (BY SIMILARITY).  
T CJEAVAGE (NONHYDROLYTIC) (BY SIMILARITY).  
Q SEQUENCE 385 AA; 41445 MW; 5A523296D0AD7C2F CRC64;

Query Match 8.4%; Score 83; DB 1; Length 385;  
Best Local Similarity 21.3%; Pred. No. 3.4;  
Matches 29; Conservative 23; Mismatches 44; Indels 40; Gaps 4;

Y 7 SNLSNLTIDASS---LNGVDKLLSAEVEKMLVQKAPN-----EG 45

b 262 TKLEJLTISGAXSEAQARKIGKISILNSPLVKRTIYGGDPNMGRLMAVGKVFDEPIPFEG 321

Y 46 IEVFGLLLYALAAARTSPKQCRADSVIFSNFSGERNVVVTEGDLKKVLDGCGAPJTRFT 105

b 322 LQIVFG---TLPVKEANPEFLKXSEYLNKNTTISLNVVLNVTISMKPWGC----- 373

Y 106 NKLRTFGRFTFEAYVD 121

b 371 -----DFTKVIK 378

RESULT 9  
PK AGST5 STANDARD; PRT; 536 AA.  
C Q80U94;  
D 28-FEB-2003 (Rel. 41, Created)  
E 28-FEB-2003 (Rel. 41, Last sequence update)  
F 28-FEB-2003 (Rel. 41, Last annotation update)  
I Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49) (PEP  
E carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).  
N PCKA OR AT00035 OR A3R\_C\_56.  
S Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
C Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
C Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
X NCBI\_TaxID=176299;  
N [1]  
P SEQUENCE FROM N.A. PubMed=11743193;  
X MEDLINE=21608550; PubMed=11743193;

A Wood D.W., Sebubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
A Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
A Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
A Chapman P., Clendenning J., Deatherage G., Gillet M., Grant C.,  
A Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
A Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon J.,  
A Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
A Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
A Chumley F., Lingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
A Nester E.W.;  
A "The genome of the natural genetic engineer Agrobacterium tumefaciens  
C58.";  
I Science 294:2317-2323(2001).  
Y [2]  
P SEQUENCE FROM N.A.  
X MEDLINE=21608551; PubMed=11743194;  
A Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
A Quorllo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,

RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
RA Wolam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,  
RA Cielo C., Slater S.;  
RT "Genome sequence of the plant pathogen and biotechnology agent  
RT Agrobacterium tumefaciens C58.";  
RL Science 294:2323-2328(2001).  
CC -!- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate  
CC + CO(2).  
CC -!- PATHWAY: Rate-limiting gluconeogenic enzyme.  
CC -!- SUSCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [ATP]  
CC family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE008978; AAL41066.1; ALT\_INIT.  
DR EMBL; AE007946; AAK85859.1; ALT\_INIT.  
DR HAMAP; MF\_00453; -; 1.  
DR InterPro; IPR001272; PEPCK\_ATP.  
DR Pfam; PF01293; PEPCK\_ATP; 1.  
DR ProDom; PD004723; PEPCK\_ATP; 1.  
DR TIGRFAMs; TIGR00224; pCKA; 1.  
DR PROSITE; PS00532; PEPCK\_ATP; 1.  
KW Gluconeogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.  
FT NP\_BIND 236 243 ATP (BY SIMILARITY).  
SQ SEQUENCE 536 AA; 57882 MW; 6F90B8F484C86A5A CRC64;

Query Match 8.4%; Score 83; DB 1; Length 536;  
Best Local Similarity 26.9%; Pred. No. 5.1;  
Matches 47; Conservative 18; Mismatches 72; Indels 38; Gaps 9;

QY 20 LNVGVKKLLSAEVEKMLVQKAPNEGIEVVVFGLL--LYALAAARTTS---PKVQRADSDV 74

b 237 LSGTGKTTLSADPARTLIGDDEHGHEGIFNFECCGYAKAKLSSEAEPEIYAA----- 291

QY 75 FSNFSGE--RNVVVTEGDLKKVLDGCGAPLRTFRNKLRTFTTEAYVDFCIAYK-HKLP 131

b 292 -TNRFGTVLENVVLDESVPDFNDN--SLTENTRS-----AYLHPFP 331

QY 132 CLNKAARELGIPAEYSVAADFLGTCPLSEL--QOSRKMFAWMYALK---TEGGV 181

b 332 NASETGIAGHPKTIMLTADAFGLVLPPIARLTPEQAWYHFLSGYTKAVAGTEKGV 386

RESULT 10

ACKA\_THETN

ID ACKA\_THETN STANDARD; PRT; 401 AA.

AC Q899V4;

DI 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Acetate kinase (EC 2.7.2.1) (Acetokinase).

GN ACKA OR TTE1481.

CS Thermoanaerobacter tengcongensis.

OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;

CC Thermoanaerobacteriaceae; Thermoanaerobacter.

OX NCBI\_TaxID=119072;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MB4 / JCM 11007;  
RX MEDLINE=21992816; PubMed=1197336;  
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
RA Chen Y., Xue Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
RT "A complete sequence of T. tengcongensis genome.";  
RL Genome Res. 12:689-700(2002).

```

-!- CATALYTIC ACTIVITY: ATP + acetate = ADP + acetyl phosphate.
-!- PATHWAY: Conversion of acetate to acetyl-CoA: first step.
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-!- SIMILARITY: Belongs to the acetylkinase family.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AE013105; RAN24700.1; -.
HAMAP; MF_00020; -.
InterPro; IPR000950; Acetate_kin.
InterPro; IPR004372; ACKA.
Pfam; PF00871; Acetate_kinase; 1.
TIGRFAMs; TIGR00016; ACKA; -.
PROSITE; PS01075; ACETATE_KINASE_1; 1.
PROSITE; PS01076; ACETATE_KINASE_2; 1.
Transferrase; Kinase; Complete proteome.
SEQUENCE 401 AA; 44361 MW; FD71C642F1E019B0 CRC64;

Query Match      8.2%; Score 81.5; DB 1; Length 401;
Best Local Similarity 19.3%; Pred. No. 5;
Matches 40; Conservative 35; Mismatches 87; Indels 45; Gaps 7;

22 GVDKLLKLAEEV---FMVLVQKGFN--EGIEVVFGLL-YALAAATTPSKVQRADSVIF 75
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
35 GINDSLTHQAEQEKVKIQREKMKHKEAIQVLEVLVDKEIGV-KMKKEIDAVGHRVWH 94
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:

76 SNSFGERNVVTGGDLKKVLQSC---APLTRTN-----KLTQRTFTIE 117
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:

95 GGEYTFDSVLIDDEVIRK-JEDCIDLAPLHPANIEG-KACQIIMPGVPMVAEVDATFHQ 153
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:

118 AYVDFCIAY-----KHLPLQNAAEIGIPAEISYLAADPLG-GCP 157
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:

154 TYPDYAYIVFPIEYVEYKHIRYRFGHTSHKYSVMRAAILGRPIELKIVTCHJNGA 213
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:

158 KLSLEQQRKMFASVYALKTEGVNVT 184
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:

214 SITAVKNGKSIDTNSGFTPLEGLAVGT 240
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:

SULT 11
N2_BORBU
LON2_BORBU STANDARD; PRT; 813 AA.
OS1558;
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
ATP-dependent protease la homolcg (EC 3.4.21.-).
BB0613.
Borrelia burgdorferi (Lyme disease spirochete).
Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;
[1]
SEQUENCE FROM N.A.
STRAIN=ATCC 35210 / B31;
MEDLINE=98065943; PubMed=9403685;
Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
van Vugt R., Palmer N., Adams K.D., Gocayne J.C., Weidman J.,
Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
Smith H.O., Venter J.C.;
"Genomic sequence of a Lyme disease spirochaete, Borrelia
burgdorferi."
Nature 390:580-586(1997).
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S16.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001162; AAC66962.1; -.
CC PIR; D70176; D70176.
CC MEROPS; S16.UPW; -.
CC TIGR; BB6613; -.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR031959; AAA_ATPase_cent.
CC InterPro; IPR003111; LON.
CC InterPro; IPR001984; Lon_endopep.
CC InterPro; IPR004815; Lon_fam.
CC Pfam; PF00004; AAA; 1.
CC Pfam; PF02190; LON; 1.
CC Pfam; PF05362; Lon_C; 1.
CC PRINTS; PR00830; ENDOLAPTASE.
CC SMART; SM00382; AAA; 1.
CC SMART; SM00464; LON; 1.
CC TIGRFAMs; TIGR00763; Lon; 1.
CC PROSITE; PS01046; LON_SER; 1.
KW Hydrolase; Serine protease; ATP-binding; Complete proteome.
FT NP_BOND 369 376 ATP (POTENTIAL).
FT ACT_SITE 719 719 BY SIMILARITY.
SQ SEQUENCE 813 AA; 92312 MW; 5129AA1498C5D0F6 CRC64;

Query Match      8.2%; Score 81.5; DB 1; Length 813;
Best Local Similarity 31.9%; Pred. No. 12;
Matches 37; Conservative 19; Mismatches 43; Indels 17; Gaps 6;

Oy 8 NLSN-LVITDASSNGVDKLLSAEVEKMLVQKGFNIEGVVFGLL-YALAAATTPSKV 66
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 475 DLSNVLFTVITANSNGSKPLLD-RYEIKVGEYSYIEKLEIAKIFLPSIKESFLDKV 533
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:

Oy 67 Q-RADSDVIFSNFGERNVVVTEG--DLKKVLDGCGAPLTRFTNKLRTFTFTTEAY 119
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 534 YRIEDDVIFNLI---RNYTMESGVRGLKRVL-----TNLIRRLVRELLVEY 577
   ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:

RESULT 12
SPSY_MOUSE
ID SPSY_MOUSE STANDARD; PRT; 366 AA.
AC P97355;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Spermine synthase (EC 2.5.1.22) (Spermidine aminopropyltransferase)
DE (SPMSY).
GN SMS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strom T.M., Francis F., Lorenz B., Boeddrich A., Econs M.J.,
RA Lehrach H., Weitingger T.;
RA "Pex gene deletions in Gy and Hyp mice provide mouse models for
RA X-linked hypophosphatemia."
RA Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA TISSUE=Diaphragm;
RA Niiranen K., Korhonen V., Janne J.;
RA "Nucleotide sequence of mouse spermidine aminopropyltransferase
RA cDNA."
RT Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]

```

SEQUENCE OF 316-366 FROM N.A.  
MEDLINE=98133937; PubMed=94673015;  
Lorenz B., Francis F., Gempel K., Boeddrich A., Joster M., Schmahl W.,  
Schmidt J., Lehrach H., Meitinger T., Strom T.M.;  
"Spermine deficiency in Gy mice caused by deletion of the spermine  
synthase gene";  
Hum. Mol. Genet. 7:541-547(1998).  
-!- CATALYTIC ACTIVITY: S-adenosylmethioninamide + spermidine = 5'-  
methylthioadenosine + spermine.  
-!- PATHWAY: BIOSYNTHESIS OF SPERMINE FROM SPERMIDINE.  
-!- SIMILARITY: BELONGS TO THE SPERMINE/SPERMINE SYNTHASES FAMILY.  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; Y09419; CAA70573.1; -  
EMBL; A5031486; AAB86631.1; -  
EMBL; AJ000093; CAA03919.1; -  
EMBL; AJ000087; CAA03918.1; -  
EMBL; AJ000088; CAA03918.1; JOINED.  
EMBL; AJ000089; CAA03918.1; JOINED.  
EMBL; AJ000090; CAA03918.1; JOINED.  
EMBL; AJ000091; CAA03918.1; JOINED.  
EMBL; AJ000092; CAA03918.1; JOINED.  
MGD; MGI:109490; Sns.  
GO; GO:0008215; P:spermine metabolism; IMP.  
InterPro: IPR001045; Sperm synthase.  
PFAM; PF01564; Sperm synthase; 1.  
PROSITE; PS01330; SPERMIDINE SYNTHASE; 1.  
Transferase.  
DOMAIN 192 232 BINDING TO DECARBOXYLATED SAM  
(POTENTIAL).  
SEQUENCE 366 AA; 41313 MW; D549F319F31C43C5 CRC64;  
Query Match 9.1%; Score 80; DB 1; Length 366;  
Best Local Similarity 22.9%; Pred. No. 6.1;  
Matches 52; Conservative 34; Mismatches 65; Indels 76; Gaps 13;  
24 DKLLSAEVEKMLVQAGNEGIEV-----FGLLYALAARTTSKVRADSDVIFSNSF 79  
128 DGRLEVEYDIEVDVDEDSYQNIKILRSKQFGNII-----ILSGVNLAESDLAYTRAI 181  
80 -----GERNVVTEGP-----LK-----KVLGGCAPLTFRTNKLRT 110  
182 MSGGKEEYTKDVLILGGGGGILCEIVLKPKNVMTWIDQVIGCKKKYR-----RT 236  
111 FGRFTT-----EAYVDFCI-----AYKHLPLQNAAEJGIPAEEDSYLAADF 152  
237 CGDVLNLRGCGCVQVLIEDCIPVWKVAKEGREFDVINDLTAVPISSTPSEDS--TWDF 294  
153 LGTCPKLS--ELQSRKMFASVVALKTEGVVN--TPVSNL--RQLGR 194  
295 LRJILSLMKVLKQCGKYF-----TOGNCVNLTAELSLYEQLGR 334  
SULT 13  
CA\_SYNY3 STANDARD; PRT; 308 AA.  
Q55497;  
15-JUL-1998 (Rel. 36, Created)  
15-JUL-1998 (Rel. 36, Last sequence update)  
16-OCT-2001 (Rel. 40, Last annotation update)  
Ornithine carbamoyltransferase, anabolic (EC 2.2.3.3) (OTCase).  
AKGF OR SL0902.  
Synecocystis sp. (strain PCC 6803).  
Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
NCBI\_Taxid=1148;  
[1]

SEQUENCE FROM N.A.  
MEDLINE=96127529; PubMed=8590279;  
Kareko C., Tanaka A., Sato S., Kotani H., Suzuki T., Miyajima N.,  
Sugitara M., Tabata S.;  
"Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
region from map positions 64% to 92% of the genome";  
J. Mol. Biol. 253:153-166(1995).  
-!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate  
+ L-citrulline.  
-!- PATHWAY: Arginine biosynthesis.  
-!- SUBCELLULAR LOCATION: Cytoplasmic.  
-!- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; D64036; BAA10847.1; -  
EMBL; PIR; S76000; S76000.  
DR HSSP; P04391; LAKM.  
DR InterPro: IPR006130; Asp/Orn Cotransf.  
DR InterPro: IPR002292; Orn carbtransf.  
DR InterPro: IPR006131; OTCase\_O.  
DR InterPro: IPR006132; OTCase\_P.  
DR Pfam; PF0185; OTCase; 1.  
DR Pfam; PF02729; OTCase N; 1.  
DR PRINTS; PR00100; AOTCASE.  
DR TIGRFAMS; TIGR00658; orn carb tr; 1.  
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
KW Arginine biosynthesis; transferase; Complete proteome.  
SQ SEQUENCE 308 AA; 33616 MW; C6CD2E4998592CFA CRC64;  
Query Match 8.0%; Score 79; DB 1; Length 308;  
Best Local Similarity 24.8%; Pred. No. 6.1;  
Matches 39; Conservative 25; Mismatches 63; Indels 30; Gaps 7;  
CY 28 LSAEVEKMLVQAGNEGIEV-----EVVFGLLYALAARTTSKVRADSDVIFSNSFGE 81  
DB 17 LTTEEMKSLQLLAADKSGVLKPHCKKILGLLFYKASTRT---RVSFYAAAYQLGGQVLD 73  
QY 82 RNVVYTE-GDLKKVLDGCAPLTFRTNKL--RTFGRTFTEAYVDFCIAYKHLPLQNAAAE 138  
DB 74 LNPSTVQVGRGEPIQDARVLDYIDILAVRTFKCTDLOTFADHA-----KYPINALSD 128  
QY 139 LGIPAEEDSYLAADFCTCPKLSLQSRKMFASVVAL 175  
DB 129 LEHP-----COILADLOTIKECFGKLEGL 152  
RESULT 14  
CPXC\_AGRU STANDARD; PRT; 422 AA.  
AC P24466;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cytochrome P450-pinF; plant-inducible (EC 1.14.-.-).  
GN CYP103 OR PINF1 OR VIRH1.  
OS Agrobacterium tumefaciens.  
OS P.asaid pt1A6.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
OX NCBI\_Taxid=358;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89213933; PubMed=2708311;  
RX Kanemoto R.H., Powell A.T., Akiyoshi D.E., Regier D.A.,  
Kerstetter R.A., Nester E.W., Hawes M.C., Gordon M.P.;  
[1]

"Nucleotide sequence and analysis of the plant-inducible locus pinf from Agrobacterium tumefaciens";  
J. Bacteriol. 171:2506-2512(1989).  
-!- FUNCTION: NOT ESSENTIAL FOR VIRULENCE, BUT MAY BE INVOLVED IN THE DETOXICATION OF PLANT PROTECTIVE AGENTS AT THE SITE OF WOUNDING.  
-!- INDUCTION: TRANSCRIPTIONALLY ACTIVATED IN THE PRESENCE OF WOUNDED PLANT TISSUE AND BY PLANT PHENOLIC COMPOUNDS, SUCH AS ACETOSYRINGONE  
-!- SIMILARITY: Belongs to the cytochrome P450 family.  
-----  
This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M19352; AAA2502.1; -;  
PIR; A32306; A32306.  
InterPro; IPR001128; Cytochrome\_P450.  
Pfam; PF00067; P450; 1.  
PRINTS; PR00385; P450.  
PROSITE; PS00086; CYTOCHROME P450; 1.  
Oxidoreductase; Monooxygenase; Electron transport; Heme; Plasmid.  
METAL 369 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
SEQUENCE 422 AA; 47519 MW; 6A9PE4AA9B7E2302 CRC64;  
-----  
Query Match 8.0%; Score 79; DB 1; Length 422;  
Best Local Similarity 23.6%; Pred. No. 8.9;  
Matches 51; Conservative 29; Mismatches 60; Indels 76; Gaps 14;  
-----  
20 LAGVDKLLSAEVEKMLVQKQAPNEGIE--VVGGLLYALARTTSPKVQRADSDVIFS 77  
70 LGTDPRTCQITELM-----NRGKAGAVDFI-----DHSMLFSN 107  
78 --SFGERNVVTGDLKYLDCAP-LTRFT-----NKLRTFGRTTEAYVDFCIAYK 127  
108 GETHGRRSGLSKAFSFRVVEALRPEIAKITECLWDLQKVDFF--NTIEMVAS----- 159  
128 HKLPQNAAEAGIPAE-----SYLAADF-GTCPLKSELQ----- 163  
160 -QLPALTIASVLGLPSEDTPFFTRLVYKVRCLSPSWRDEFEETIEASAIQLQDVRSVI 218  
164 --QSRKM---FASMYALK--EGGVNTPVSNLRQ- 192  
219 ADSGRMRDPLSRY-LKAVREAGTL-SPIEIMQL 252  
-----  
RESULT 15  
PCK\_FUSNN STANDARD; PRT; 527 AA.  
-----  
Q8REI2;  
28-FEB-2003 (Rel. 41, Created)  
28-FEB-2003 (Rel. 41, Last sequence update)  
28-FEB-2003 (Rel. 41, Last annotation update)  
Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49) (PEP carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).  
PCKA OR FN1120.  
Fusobacterium nucleatum (subsp. nucleatum).  
Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;  
Fusobacterium.  
NCBI\_TaxID=76856;  
[1]  
-----  
SEQUENCE FROM N. A.  
STRAIN=ATCC 25586;  
MEDLINE=21866394; PubMed=11889.09;  
Kapatal V., Andersson I., Ivanova N., Reznik G., Los T., Lykidis A., Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L., Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A., Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R., Fomstein X., Kyrpides N., Overbeek R.;  
"Genome sequence and analysis of the oral bacterium Fusobacterium

nucleatum strain ATCC 25586";  
J. Bacteriol. 184:2005-2018(2002).  
-!- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate + CO(2).  
-!- PATHWAY: Rate-limiting gluconeogenic enzyme.  
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
-!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase (ATP) family.  
-----  
This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AE010616; AAL95316.1; -;  
HAMAP; MF 00453; -; 1.  
InterPro; IPR001272; PEPCK\_ATP.  
Pfam; PF01293; PEPCK\_ATP; 1.  
ProDom; PD004723; PEPCK\_ATP; 1.  
TIGRFAMs; TIGR00224; pckA; 1.  
PROSITE; PS00532; PEPCK\_ATP; 1.  
KW Gluconeogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.  
NP\_BIND 230 237 ATP (BY SIMILARITY).  
FT NP\_BIND 230 237 ATP (BY SIMILARITY).  
SQ SEQUENCE 527 AA; 59055 MW; 275849PDP254AC01 CRC64;  
-----  
Query Match 8.0%; Score 79; DB 1; Length 527;  
Best Local Similarity 26.6%; Pred. No. 12;  
Matches 46; Conservative 26; Mismatches 67; Indels 34; Gaps 10;  
-----  
QY 20 LAGVDKLLSAEVEKMLV---QKQAPNEGIEVVGGLLYALARTTSPKVQRADSDVIFS 76  
Db 231 LSGTGRKTTLSADPNRKLGDDEHGWCDKGIHFNEG---GCYAKCINLK-ESESEPIYRA 285  
QY 77 NSFGE--RNVTVEGGLKKVLCGCAPLTRFNKLRTFGRTT-EAYVDFCIAYKHLPLQL 133  
Db 286 IKGSLVENVVD-----PITR---KIQVEDASITENTRUGYPIDY---TPNA 327  
QY 134 NAAAEGLG:PAEDSYLAADFLGTCPKLSELQSSRN--FASMYALK---TEGGV 181  
Db 328 ELSGVGGIPKWIFLTADSPGLVPPISRLSQEAAVHFVTGFTAKLAGTELGV 380  
-----  
Search completed: November 7, 2003, 14:49:01;  
Job time : 18 secs

GenCore version: 5.1.6  
 Copyright (c) 1993 - 2003 CompuGen Ltd.  
 November 7, 2003, 14:46:13 : Search time 34 Seconds  
 (without alignments)  
 1502.777 Million cell updates/sec

US-09-613-486-15  
 effect score: 991  
 sequence: 1 MELMSDSNLSNLVITDASSL.....GGVNVTPVSNLRQJGRREV 198

scoring table: BLCSJM62  
 Gapop 10.0 , Gapext C.5  
 searched: 830525 seqs, 258052604 residues  
 total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Maximum Match 100%  
 Listing first 45 summaries

Database: SPTREMBL\_23:  
 1: sp\_archaea:  
 2: sp\_bacteria:  
 3: sp\_fungi:  
 4: sp\_human:  
 5: sp\_invertebrate:  
 6: sp\_mammal:  
 7: sp\_rhiz:  
 8: sp\_organelle:  
 9: sp\_phage:  
 10: sp\_plant:  
 11: sp\_protist:  
 12: sp\_virus:  
 13: sp\_vertebrate:  
 14: sp\_unclassified:  
 15: sp\_virus:  
 16: sp\_bacterioph:  
 17: sp\_archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

sult No.	Score	Query Match	Length	ID	Description
1	991	100.0	198	12	071213 grapevine 1
2	984	99.3	198	12	039857 grapevine 1
3	921	92.9	198	12	088223 grapevine 1
4	297	30.0	204	12	090629 sugar beet
5	296	29.9	204	12	096664 sugar beet
6	295	29.8	204	12	089910 sugar beet
7	293	29.6	204	12	031724 sugar beet
8	286	28.9	204	12	008531 sugar beet
9	266.5	26.9	215	12	085858 beet yellow
10	130.5	13.2	223	12	080971 citrus tris
11	129.5	13.1	223	12	090629 citrus tris
12	125.5	12.7	222	12	093070 citrus tris
13	125.5	12.7	223	12	093070 citrus tris
14	125.5	12.7	223	12	090629 citrus tris
15	125	12.6	223	12	090629 citrus tris
16	124.5	12.6	223	12	090629 citrus tris

17	123.5	12.5	223	12	09E7M2	citrus tris
18	123.5	12.5	223	12	09E7M4	citrus tris
19	123.5	12.5	223	12	09FXT0	citrus tris
20	123.5	12.5	223	12	09QEG3	citrus tris
21	123	12.4	212	12	09DWM1	citrus tris
22	123	12.4	223	12	09E7M3	citrus tris
23	120.5	12.2	195	12	08V1P7	citrus tris
24	120.5	12.2	222	12	0908N9	citrus tris
25	120.5	12.2	223	12	09QEG1	citrus tris
26	120.5	12.2	223	12	089948	citrus tris
27	120.5	12.2	223	12	0909P1	citrus tris
28	120.5	12.2	223	12	09WID9	citrus tris
29	120.5	12.2	223	12	091FW4	citrus tris
30	120.5	12.2	223	12	096192	citrus tris
31	120.5	12.2	223	12	09FXT3	citrus tris
32	118.5	12.0	223	12	08FYU9	citrus tris
33	117.5	11.9	223	12	09FXS9	citrus tris
34	117.5	11.9	223	12	0991G6	citrus tris
35	116	11.7	223	12	09QEG6	citrus tris
36	115.5	11.7	195	12	08V1P5	citrus tris
37	115.5	11.7	223	12	08JYU0	citrus tris
38	115	11.6	223	12	09FXT1	citrus tris
39	114.5	11.6	223	12	09D7F8	citrus tris
40	114.5	11.6	223	12	010474	citrus tris
41	114.5	11.6	223	12	08B3R9	citrus tris
42	113.5	11.5	194	12	08V1P6	citrus tris
43	111.5	11.3	223	12	08JYU8	citrus tris
44	110.5	11.2	223	12	09FXT2	citrus tris
45	109	11.0	203	12	099AT6	citrus tris

## ALIGNMENTS

RESULT 1  
 071213 PRELIMINARY: PRT; 198 AA.  
 AC 071213  
 DT 071213 (T-EMBLrel. 07, Created)  
 DC 071213 (T-EMBLrel. 07, Last sequence update)  
 D7 071213 (T-EMBLrel. 17, Last annotation update)  
 DE 22 kDa coat protein.  
 OS Grapevine leafroll-associated virus 2.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
 OC Closterovirus.  
 OX NCBI\_TaxID=64003;  
 RN [1]  
 RX SEQUENCE FROM N.A.  
 RP MEDLINE=98264507; PubMed=9603345;  
 RA Zhu H.Y., Ling K.S., Gosczyrski J.E., McPerson J.R., Gonsalves D.;  
 RT "Nucleotide Sequence and Genome Organization of Grapevine Leafroll-  
 RT Associated Virus-2 are Similar to Beet Yellow Virus, the  
 RT Closterovirus Type Member."  
 RL J. Gen. Virol. 79:1289-1298(1998).  
 DR EMBL; AF039204; AAC40861.1; -;  
 DR InterPro; IPR002679; Closter\_coat.  
 DR Pfam; PF01785; Closter\_coat; 1.  
 KW Coat protein.  
 SQ SEQUENCE 198 AA; 21660 MW; 8ED5EF2EFIDPB03 CRC64;  
 Query Match 100.0%; Score 991; DB 12; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-89;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MELMSDSNLSNLVITDASS:NGVDKLLSAEVEKMLVQKGNEGTEVVFGLLVALAAR 60  
 Ds 1 MELMSDSNLSNLVITDASS:NGVDKLLSAEVEKMLVQKGNEGTEVVFGLLVALAAR 60  
 Cy 61 TTSPKQVQADSDVIFNSFGERNVVVTEGDLKVLGDGCAPLTFRTNKLRTFGRTTEAYV 120  
 Ds 61 TTSPKQVQADSDVIFNSFGERNVVVTEGDLKVLGDGCAPLTFRTNKLRTFGRTTEAYV 120  
 Cy 121 DFCIAYKHKLPLQNAALAEGLGPAEDSYLAADFLGTCTCKLSEIQQRKMFASMYALKTEGG 190

121 DFCAYKHKLPOLNAAELG:PAEDSYLAADFLGTCPKLSELOQSRKVFASMYALKTEGG 180  
181 VVNTPVSNLRQLGRREV 198  
181 VVNTPVSNLRQLGRREV 198

## SULT 2

9857 O39857 PRELIMINARY: PRT; 198 AA.  
O39857;  
01-JAN-1998 (TReMBLrel. 05, Created)  
01-JAN-1998 (TReMBLrel. 05, Last sequence update);  
01-DEC-2001 (TReMBLrel. 19, Last annotation update);  
Coat protein.  
Grapevine leafroll-associated virus 2.  
Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
Closterovirus.  
NCBI\_TaxID=64003;  
[1]  
SEQUENCE FROM N.A.  
Abou-Gharem N.;  
"The nucleotide sequence of the 3'terminal region of grapevine  
leafroll associated closterovirus 2.";  
Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
EMBL: Y14131; CAA74566.1; -;  
InterPro: IPR002679; Closter\_coat.  
Pfam: PF01785; Closter\_coat; 1.  
SEQUENCE 198 AA; 21630 MW; 15D631E4088D3F01 CRC64;

Query Match 99.3%; Score 984; DB 12; Length 198;

Best Local Similarity 99.5%; Pred. No. 3.7e-88;  
Matches 197; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 MELMSDSNLSNLVITDASSLNGVDKLLSAEVEKMLVOKGAPNEGIEVVFGLLLYALAAR 60  
1 MELMSDSNLSNLVITDASSLNGVDKLLSAEVEKMLVOKGAPNEGIEVVFGLLLYALAAR 60  
61 TTSPKVQRADSDVIFSNFGERNVVVTGGDKKVLGGCAPLRTFTNKLRTFGRTTEAYV 120  
61 TTSPKVQRADSDVIFSNFGERNVVVTGGDKKVLGGCAPLRTFTNKLRTFGRTTEAYV 120  
121 DFCIAYKHKLPOLNAAELG:PAEDSYLAADFLGTCPKLSELOQSRKVFASMYALKTEGG 180  
121 DFCIAYKHKLPOLNAAELG:PAEDSYLAADFLGTCPKLSELOQSRKVFASMYALKTEGG 180  
181 VVNTPVSNLRQLGRREV 198  
181 VVNTPVSNLRQLGRREV 198

## SULT 3

9857 O39857 PRELIMINARY: PRT; 198 AA.  
O39857;  
01-MAR-2003 (TReMBLrel. 23, Created)  
01-MAR-2003 (TReMBLrel. 23, Last sequence update);  
01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
Coat protein.  
Grapevine rootstock stem lesion associated virus.  
Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
Closterovirus.  
NCBI\_TaxID=167634;  
[1]  
SEQUENCE FROM N.A.  
Zhang Y., Rowhani A.;  
"Nucleotide sequence of grapevine rootstock stem lesion associated  
virus.";  
Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
EMBL: AF314061; AAN63472.1; -;  
SEQUENCE 198 AA; 21620 MW; 370E95A9369734A9 CRC64;

Query Match 92.9%; Score 921; DB 12; Length 198;  
Best Local Similarity 90.9%; Pred. No. 5.4e-82;  
Matches 180; Conservative 10; Mismatches 8; Indels 0; Gaps 0;  
QY 1 MELMSDSNLSNLVITDASSLNGVDKLLSAEVEKMLVOKGAPNEGIEVVFGLLLYALAAR 60  
DB 1 MELMSDDNLSGLVITDASSLNGVDKLLSAEVEKMLVOKGAPSQGIEVVFGLLLYALAAR 60  
QY 61 TTSPKVQRADSDVIFSNFGERNVVVTGGDKKVLGGCAPLRTFTNKLRTFGRTTEAYV 120  
DB 61 TTSPKVQRADSDVIFSNFGERNVVVTGGDKKVLGGCAPLRTFTNKLRTFGRTTEAYV 120  
QY 121 DFCIAYKHKLPOLNAAELG:PAEDSYLAADFLGTCPKLSELOQSRKVFASMYALKTEGG 180  
DB 121 DFCVAYKHKLPOLNAAELG:PAEDSYLAADFLGACPKLSELOQSRKVFASMYALKTEGG 180  
QY 181 VVNTPVSNLRQLGRREV 198  
DB 181 VVNTPVSNLRQLGRREV 198

## RESULT 4

Q3Q629  
ID Q3Q629 PRELIMINARY: PRT; 204 AA.  
AC Q3Q629;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2003 (TReMBLrel. 13, Last annotation update);  
DE Major capsid protein.  
DE Sugar beet yellow virus (SBYV).  
CC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
CC Closterovirus.  
CX NCBI\_TaxID=12161;  
RN [1]  
SEQUENCE FROM N.A.  
STRAIN=BYV-4;  
RX MEDLINE=20079557; PubMed=10611288;  
RX Peremyslov V.V., Hagiwara Y., Dolja V.V.;  
RT "HSP70 homolog functions in cell-to-cell movement of a plant virus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:14771-14776(1999).  
RN [2]  
SEQUENCE FROM N.A.  
STRAIN=BYV-4;  
RA Peremyslov V.V., Hagiwara Y., Alzhanova D., Dolja V.V.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF190581; RAF14305.1; -;  
DR InterPro: IPR002679; Closter\_coat.  
DR Pfam: PF31785; Closter\_coat; 1.  
SQ SEQUENCE 204 AA; 22419 MW; FD6337E1D2490D33 CRC64;

Query Match 30.0%; Score 297; DB 12; Length 204;  
Best Local Similarity 37.6%; Pred. No. 6.4e-21;  
Matches 71; Conservative 32; Mismatches 84; Indels 2; Gaps 2;  
QY 7 SNLSNLVITDASSLNGVDKLLSAEVEKMLVOKGAPNEGIEVVFGLLLYALAAR:TSRKV 66  
DB 11 ATFENVSLADOTCLMGEDCDKLRDFECLKLVGPEDKLGALGLCLYSCATIGTSNKV 70  
QY 67 QRADSDVIFSNFGE-RNVVVTGGDKKVLGGCAPLRTFTNKLRTFGRTTEAYVDFCIA 125  
DB 71 SVQSTTF:KASFGSGKELFLTHGELRSFLDSQKLEGGKFNKLCFCFTFKDYISFAKE 130  
QY 126 YXHKLPOLNAAELG:PAEDSYLAADFLGTCPKLSELOQSRKVFASMYALKTEGGVWNTP 185  
DB 131 YGRLLPPIARANPHGLPAEDHYLAADF:STSTELTDLOQGRLLARENATHTFS-SESP 189  
QY 186 VSNLRQLGR 194  
DB 190 VTSLKQLGR 198

## RESULT 5

Q36664



```

RP      SEQUENCE FROM N.A.
RC      STRAIN=German;
RA      Agrasovsky A.A., Koonin E.Y., Boyko V.P., Maiss E., Lurina N.A.,
RA      Arabekov J.G.;
RT      "Beet yellows closterovirus: complete genome structure and
RT      identification of a leader papain-like thiol protease";
RL      Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
DR      EMBL; M59452; AAA72955.1; -
DR      EMBL; X73475; CAA51861.1; -
DR      InterPro: IPR002679; Closter_coat.
DR      Pfam: PF01785; Closter_coat; 1.
SQ      SEQUENCE 204 AA; 22388 MW; 92B2A4DCA77543D9 CRC64;

Query Match          29.6%; Score 293; DB 12; Length 204;
Best Local Similarity 37.6%; Pred. No. 1.6e-20;
Matches 71; Conservative 31; Mismatches 85; Indels 2; Gaps 2;

QY      7  SNLSLVITDASSLNGVDKLLSAEYKMLVCKGAPNEGIEVVFGLLLYALAAARTTSPKV 66
Db      11  ATENYSLADQTCLEHGDCDCKURKQNFCECLKLGVPEDKLGLALGLCLYSCATIGTSNKV 70

QY      67  QRADSOVIESNSF-GRNVVVVTGGDLKKVLDGCAPLRTFTNKLRTFGRTFTAYYDFCIA 125
Db      71  SVQSTFTFKASFGGKKEFLTHGELRSLFSDSKLLEGKPNKJRCFCRTFQKDYISFAKE 130

QY      126  YKHKLQPLNAAAEGLIPADSDSYLAADFLGTCPKLSELOQSRKXFASMYALKTEGGVVNT? 185
Db      132  YGRGLPPIAARNHGLPAEDHYLAADFISTSTELTDQGRLLARENATHTEFS-SESP 189

QY      186  VSNLRQZGR 194

```

11:11:11  
190 VTSKQLGR 198

SULT 8

8531

Q08531 PRELIMINARY; PRT; 204 AA.

01-NOV-1996 (TrEMBLrel. 01, Created)

01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

Coat protein.

Sugar beet yellow virus (SBYV).

Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;

Closterovirus.

NCBI\_TaxID=1216;

[1]

SEQUENCE FROM N.A.

MEDLINE=91116305; PubMed=1990061;

Agranovsky A.A., Boyko V.P., Karasev A.V., Lunina N.A., Koonin E.V.,

Dolja V.V.

"Nucleotide sequence of the 3'-terminal half of beet yellows

closterovirus RNA genome: unique arrangement of eight virus genes."

J. Gen. Virol. 72:15-23(1991).

[2]

SEQUENCE FROM N.A.

STRAIN=UKRAINIAN;

MEDLINE=94082464; PubMed=8259666;

Agranovsky A.A., Koonin E.V., Boyko V.P., Mais E., Froetschl R.,

Lunina N.A., Atabekov J.G.

"Beet yellows closterovirus: complete genome structure and

identification of a leader papain-like thiol protease."

Virology 198:311-324(1994).

EMBL: X53462; CAA37554.1; -

EMBL: X73476; CAA51868.1; -

InterPro: IPR002679; Closter\_coat.

Pfam: PF01785; Closter\_coat; 1.

Coat protein.

SEQUENCE 204 AA; 22356 MW; 538E8E5811831486 CRC64;

Query Match

28.9%; Score 286; DB 12; Length 204;

Best Local Similarity 37.6%; Pred. No. 7.7e-20;

Matches 71; Conservative 29; Mismatches 87; Indels 2; Gaps 2;

7 SLSLVITDASSLNGVKKL-SAEVEKMLVOKGAPNEG-EVVFGLLYALAAARTSPKV 66

11 ATENVSLAQITLHGDCCKLRNFEECLKGVPEDNGLGALGLCLYSCATIGTSNKV 70

67 QRADSDVIFNSF-GERNVVVTEGLKKVLDGCAPLRTFNKLRITFGRTTEAYVDFCIA 125

71 NVQPTSTFIKASFGGKELY-THGELNSFLGSQKLLSGKPKLRCFCRTFKDYISLRKE 130

126 YKHKLPGLNAAELGIPAEYSYLAADFLGTCPLKSELOQSRKMFASMYALKTEGGVWNTP 185

131 YRGKLPPIARNRHGLPAEDHYLAADFIPTSTELTDLQSRLLARENATHTFEFS-SESP 189

186 VSNLRQLGR 194

190 VTSKQLGR 198

SULT 9

5858

Q65858 PRELIMINARY; PRT; 215 AA.

01-NOV-1996 (TrEMBLrel. 01, Created)

01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

Coat protein.

Beet yellow stunt virus.

Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;

Closterovirus.

NCBI\_TaxID=35290;

RN

RP SEQUENCE FROM N.A.

RX MEDLINE=96266429; PubMed=8661428;

RA Karasev A.V., Nikolaeva O.V., Mushegian A.R., Lee R.F., Dawson W.O.;

RT "Organization of the 3'-terminal half of beet yellow stunt virus

genome and implications for the evolution of closteroviruses."

R. Virology 221:199-207(1996).

RN

RP SEQUENCE FROM N.A.

RA Karasev A.V., Nikolaeva O.V., Lee R.F., Wisler G.C., Duffus J.E.,

RA Dawson W.O.;

RT "Characterization of the beet yellow stunt virus coat protein gene."

R. Phytopathology 88:1040-1045(1998).

DR EMBL: U51931; AAC55665.1; -

DR InterPro: IPR002679; Closter\_coat.

DR InterPro: IPR002048; EF-hand.

DR Pfam: PF01785; Closter\_coat; 1.

DR PROSITE: PS00018; EF\_HAND; 1.

SQ SEQUENCE 215 AA; 23696 MW; 91185835840806FA CRC64;

Query Match

26.9%; Score 266.5; DB 12; Length 215;

Best Local Similarity 36.1%; Pred. No. 6.7e-18;

Matches 66; Conservative 22; Mismatches 94; Indels 1; Gaps 1;

QY

11 NLVITDASSLNGVKKL-SAEVEKMLVOKGAPNEG-EVVFGLLYALAAARTSPKVQCAD 70

Db

30 NPARASATCLNGENKKLFEFSVRVKTQDVTESGIPTTLGLTLYALATLSTSSKIDED 89

QY

71 SDVIFNSFGERNVVVTEGLKKVLDGCAPLRTFNKLRITFGRTTEAYVDFCIAVYKHL 130

Db

90 KTPVLSAKIDAVNVITVEDIKNFVNSLTLLKNYKLVFARTFEEYLVRCYKHL 149

QY

131 POLNAAAEELGIPAEYSYLAADFLGTCPLKSELOQSRKMFASMYALKTEGGVWNTPVSNLR 190

Db

150 PNIRANKHGIPADYSYLAADFVQTSNLLKEHA-VLLEGRNAATASSGTTRESAVNLK 208

QY

191 QQG 193

Db

209 YJG 211

RESULT 10

Q8JYVI

ID Q8JYVI PRELIMINARY; PRT; 223 AA.

AC Q8JYVI

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Coat protein.

OS Citrus tristeza virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;

OC Closterovirus.

OX NCBI\_TaxID=12162;

RN

RP SEQUENCE FROM N.A.

RA Roy A., Ramachandran P., Brlansky R.H.;

RT "Molecular characterization of Indian Citrus tristeza virus

isolates."

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF501867; AAM22216.1; -

DR InterPro: IPR002679; Closter\_coat.

DR Pfam: PF01785; Closter\_coat; 1.

SQ SEQUENCE 223 AA; 24903 MW; 9D31F82AD158B649 CRC64;

Query Match

13.2%; Score 130.5; DB 12; Length 223;

Best Local Similarity 27.7%; Pred. No. 0.00014;

Matches 61; Conservative 27; Mismatches 75; Indels 57; Gaps 11;

QY

1 MELMSDSNLSNLVITDASSLNGVKKLSAEVEKMLVOKG-----APNEGIEVVFGLLYA 56

Db

31 MNLH:DPIT--IAMDVRQLGTQQNAALNRL--FLTLKGYKPNLPDSDKDFHAMMLYR 86

QY

57 LAARTS-----PKVQAD---SDVIF-SNSFGERNVVVTEGLKKVLD 96

```

106 87 LAVKSSLSQDDTTGITYTRGVEVDLSKLTWDFVNSKIGNR----- 132
107 97 GCAPLTFRTNKLRTGRTTAYVDFCIAYKHKLPLQNAAEELGIPAECSYLAADFLGTC 156
108 133 -----TNALRVGRTNDALYLAF-C-RQNRNLSYGGRRPLDAGIPAGVHYLCADFL-TG 182
109 157 PKLSELOQSRKFPASMYALKTEGG--VNTFVSNLRQLGR 194
110 183 AGLTDLCAVYIOAKEQLKKRGAEW---VTNVRQLGK 219
111
112 RESULT 11
113 90EG4 PRELIMINARY; PRT; 223 AA.
114 Q90EG4;
115 Q90EG4;
116 01-MAY-2000 (TrEMBLrel. 13, Created;
117 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
118 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
119 Coat protein.
120 Citrus tristeza virus.
121 Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
122 Closterovirus.
123 NCBI_TaxID=12162;
124 {1}
125 SEQUENCE FROM N.A.
126 STRAIN=25-20;
127 Sequiera Z., Nolasco G.;
128 "Bacterial expressed coat protein for the development of a single
129 antibody for routine detection of Citrus Tristeza Virus in Portugal.";
130 Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
131 {2}
132 SEQUENCE FROM N.A.
133 STRAIN=25-20;
134 Bonacalza B., Febres V., Niblett C.J., Nolasco G.;
135 "Biological characterization of Citrus Tristeza Virus isolates from
136 Portugal.";
137 Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
138 {3}
139 EMBL; AF184115; AAD56568.1; -.
140 InterPro: IPR002679; Closter.coat.
141 Pfam: PF01785; Closter.coat.1.
142 SEQUENCE 223 AA; 24930 MW; 05C4C6BD17C2D7F0 CRC64;
143
144 Query Match 13.1%; Score 129.5; DB 12; Length 223;
145 Best Local Similarity 26.3%; Pred. No. 0.0018;
146 Matches 55; Conservative 31; Mismatches 68; Indels 55; Gaps 10;
147
148 12 LVITDASSLNGVDKLLSAEVEKMLVQKG----APNEGIEVVFGLLYALAAARTTS---- 63
149 40 ITMNEVRLSTQQAALNRDL--FLALRGKYPNLPDKDKDFHIAVLMYRLAVKSSLSQSD 97
150 64 -----PKVQRAD---SDVIF-SNSFGERNVVTGDLKKVLDGCAPLTRFTNK 107
151 98 DDTTGITYTRGVEVDLSKLTWDFVNSKIGNR-----TNA 135
152 108 LRTFGRTTAYVDFCIAYKHKLPLQNAAEELGIPAECSYLAADFLGTCPKLSELOQSRK 167
153 136 LRWVGRTNDALYLAF-C-RQNRNLSYAGRPDLADGIPAGVHYLCADFL-TGAGLTDLCAVY 193
154 168 MFASMYALKTEGG--VNTFVSNLRQLGR 194
155 194 IOAKEQLKKRGAEW---VTNVRQLGK 219
156
157 Query Match 12.7%; Score 125.5; DB 12; Length 223;
158 Best Local Similarity 26.0%; Pred. No. 0.00044;
159 Matches 56; Conservative 29; Mismatches 69; Indels 55; Gaps 10;
160
161 12 LVITDASSLNGVDKLLSAEVEKMLVQKG----APNEGIEVVFGLLYALAAARTTS---- 63
162 39 IAKNDVRLSTQQAALNRDL--FUTLKGKYPNLPDKDKDFHIAVLMYRLAVKSSLSQSD 96
163 64 -----PKVQRAD---SDVIF-SNSFGERNVVTGDLKKVLDGCAPLTRFTNK 107
164 97 DDTTGITYTRGVEVDLSKLTWDFVNSKIGNR-----TNA 134
165 108 LRTFGRTTAYVDFCIAYKHKLPLQNAAEELGIPAECSYLAADFLGTCPKLSELOQSRK 167
166 135 LRWVGRTNDALYLAF-C-RQNRNLSYAGRPDLADGIPAGVHYLCADFL-TGAGLTDLCAVY 192
167 168 MFASMYALKTEGG--VNTFVSNLRQLGR 194
168 193 IOAKEQLKKRGAEW---VTNVRQLGK 218
169
170 RESULT 13
171 Q990P2 PRELIMINARY; PRT; 223 AA.
172 Q990P2;
173 01-JUN-2001 (TrEMBLrel. 17, Created)
174 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
175 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
176 DE Coat protein.
177 Citrus tristeza virus.
178 Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
179 Closterovirus.
180 NCBI_TaxID=12162;
181 {1}
182 SEQUENCE FROM N.A.
183 STRAIN=TAM11;
184 Herrera-Isidron L., Villarreal-Garcia L.A., Rivera-Bustamante R.,
185 Martinez-Soriano J.P.;
186 "Molecular Analysis of Citrus tristeza virus from Mexico.";
187 Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
188 EMBL; AF342890; AAK27476.1; -.
189 InterPro: IPR002679; Closter.coat.
190 Pfam: PF01785; Closter.coat.1.
191 SEQUENCE 223 AA; 24868 MW; 6EAC742D2212949 CRC64;
192
193 Query Match 12.7%; Score 125.5; DB 12; Length 223;
194 Best Local Similarity 27.0%; Pred. No. 0.00044;
195 Matches 60; Conservative 27; Mismatches 74; Indels 61; Gaps 11;
196
197 1 MELMSDNLNLSNVITDASSLNGVDKLLSAEVEKMLVQKG----APNEGIEVVFGLLYA 56
198 31 MNJHIDFTL--TAMNDVRLSTQQAALNRDL--FUTLKGKYPNLPDKDKDFHIAVLMY 86
199 57 LAARTTSPKVQRAD-----SDVIF-SNSFGERNVVTGDLKKV 94
200 87 LAVKSS--LQSDDDTTGITYTRGVEVDLPDKLTWDFVNSKIGNR----- 132
201 95 LDGCAPLTRFTNKLRTGRTTAYVDFCIAYKHKLPLQNAAEELGIPAECSYLAADFLG 154
```

```

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
OC Closterovirus.
OX NCBI_TaxID=12162;
RN {1}
RC SEQUENCE FROM N.A.
RC STRAIN=BC15-1;
RA Herrera-Isidron L., Villarreal-Garcia L.A., Rivera-Bustamante R.,
RA Martinez-Soriano J.P.;
RA "Molecular Analysis of Citrus tristeza virus from Mexico.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF342894; AAK27480.1; -.
DR InterPro: IPR002679; Closter.coat.
DR Pfam: PF01785; Closter.coat.1.
DR SEQUENCE 222 AA; 24741 MW; 5BFA27BCB4B08DDE CRC64;
SQ

```

```

Query Match 12.7%; Score 125.5; DB 12; Length 222;
Best Local Similarity 26.0%; Pred. No. 0.00044;
Matches 56; Conservative 29; Mismatches 69; Indels 55; Gaps 10;
12 LVITDASSLNGVDKLLSAEVEKMLVQKG----APNEGIEVVFGLLYALAAARTTS---- 63
39 IAKNDVRLSTQQAALNRDL--FUTLKGKYPNLPDKDKDFHIAVLMYRLAVKSSLSQSD 96
64 -----PKVQRAD---SDVIF-SNSFGERNVVTGDLKKVLDGCAPLTRFTNK 107
97 DDTTGITYTRGVEVDLSKLTWDFVNSKIGNR-----TNA 134
108 LRTFGRTTAYVDFCIAYKHKLPLQNAAEELGIPAECSYLAADFLGTCPKLSELOQSRK 167
135 LRWVGRTNDALYLAF-C-RQNRNLSYAGRPDLADGIPAGVHYLCADFL-TGAGLTDLCAVY 192
168 MFASMYALKTEGG--VNTFVSNLRQLGR 194
193 IOAKEQLKKRGAEW---VTNVRQLGK 218

```

UT 15  
T3  
Q8B3T3  
Q8B3T3; PRELIMINARY; PRT; 223 AA.  
01-MAR-2003 (TRENBLrel. 23, Created)  
01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
25 kDa major coat protein.  
Citrus tristezza virus (isolate T36) (CTV).  
Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae.  
Closterovirus.  
NCBI\_TaxID=31712;  
[1]  
[1]

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - nucleic search, using frame\_plus\_p2n model

run on: November 7, 2003, 20:42:48 ; Search time 2044 Seconds  
(without alignments)  
2354.347 Million cell updates/sec

title: US-09-613-486-15  
effect score: 991  
sequence: 1 MEIMSDSNLSNLVITASSL.....GGVNTPTVSNLRQLGRREV 198

scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

searched: 22781392 seqs, 1252238056 residues

total number of hits satisfying chosen parameters: 45562784

minimum DB seq length: 0

maximum DB seq length: 2000000000

post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

>command line parameters:

>>>DBL=frame\_p2n.model -DBV=nlh  
>>>=/cgr2:/usrpro.spool/US09613486/runat\_07112003\_120411\_27105/app\_query.fasta\_1.391  
>>>B=EST -QFMT=fastap -SUFFIX=txt -MINMATCH=0.1 -LOOPCL=C -LOOPEXT=0  
>>>NITS=bits -START=1 -END=1 -MATRIX=blomsm62 -TRANS=human43.cdi -LIST=45  
>>>OCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
>>>UTFMT=ptc -NCRW=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
>>>SFR=US09613486 @CGN\_1.1 2810 @runat\_07112003\_120411\_27105 -NCPU=6 -ICPU=3  
>>>O.WMAP -LARGEQUERY -NEG\_SCORES=C -WAIT -DSPLOCK=100 -LONGLOG  
>>>EV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
>>>GAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

tabase : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmd:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estto:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fur:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	88.5	8.9	706	13	BQ986943	BQ986943 QGF10N21.
2	88.5	8.3	748	28	BZ051926	BZ051926 jns57n07.
3	87	8.8	477	12	BW132093	BW132093 TGE57ybo
4	85.5	8.6	562	13	BQ459495	BQ459495 QA08017r
5	85.5	9.6	666	13	BUC07723	BUC07723 QGH5E15.Y
6	85.5	8.6	671	12	BJ433495	BJ433495 BJ433495
7	85.5	8.6	704	29	B2434184	B2434184 BONPR91TF
8	85.5	8.6	713	9	AU034876	AU034876 AU034876
9	85.5	8.6	751	12	BJ376355	BJ376355 BJ376355
10	85	8.6	654	13	BQ916284	BQ916284 QHB17604.
11	84.5	8.5	610	13	BQ848646	BQ848646 QGA7M01.Y
12	84.5	8.5	611	13	BQ984758	BQ984758 QGE5A05.Y
13	84.5	8.5	632	10	BG526441	BG526441 61-95 Ste
14	84.5	8.5	648	10	BG523153	BG523153 29-48 Ste
15	84.5	8.5	695	13	BQ993014	BQ993014 QGF27R14.
16	84	8.5	589	13	BQ82386	BQ82386 QGB16E23.
17	84	8.5	600	10	BF598403	BF598403 sv17B66.Y
18	84	8.5	732	28	BH976504	BH976504 cdh67E01.
19	84	8.5	1098	10	BG247808	BG247808 602359515
20	83.5	8.4	633	28	BH096256	BH096256 RPCI-24-2
21	83.5	8.4	644	12	BJ435395	BJ435395 BJ435395
22	83.5	8.4	644	12	BJ435395	BJ435395 BJ435395
23	83.5	8.4	704	14	C921323	C921323 C921323
24	83	8.4	600	14	C8343251	C8343251 CA32EN030
25	83	8.4	511	10	BQ421690	BQ421690 60249789
26	82.5	8.3	661	28	AQ349879	AQ349879 Sheared D
27	82.5	8.3	762	14	CB569015	CB569015 AGENCOURT
28	82	8.3	607	12	B1773596	B1773596 rs31h33.Y
29	82	8.3	627	13	BQ26578	BQ26578 QHG17E05.
30	82	8.3	750	13	BQ481813	BQ481813 PV_GEA012
31	82	8.3	528	13	BU410316	BU410316 603158539
32	82	8.3	523	14	CA582924	CA582924 ES-002599
33	82	8.3	1028	13	BU202935	BU202935 603102659
34	91.5	8.2	428	10	BG544734	BG544734 E2700 Chi
35	81.5	8.2	560	29	CG155340	CG155340 CSJ-K34.1
36	81.5	8.2	606	28	AQ079289	AQ079289 RPCI-23-4
37	81	8.2	513	28	AQ951071	AQ951071 Sheared D
38	91	8.2	727	28	BH929686	BH929686 odh98c12.
39	91	8.2	807	29	B2577160	B2577160 rsn2_5276
40	80.5	8.1	836	12	BI085103	BI085103 602870643
41	80	8.1	432	13	BU765335	BU765335 sas16f01.
42	80	8.1	717	13	BQ969618	BQ969618 QHB39B11.
43	80	8.1	751	14	CA316784	CA316784 UI-M-FWC-
44	80	8.1	868	28	AZ691220	AZ691220 ENTJ241TF
45	80	8.1	999	12	B1199349	B1199349 602759615

# ALIGNMENTS

RESULT 1  
BQ986943  
LOCUS  
DEFINITION BQ986943 706 bp mRNA linear EST 21-AUG-2002  
QGF10N21.YG.ab1 QG\_EFGHJ lettuce serriola Lactuca sativa cDNA clone  
QGF10N21, mRNA sequence.  
ACCESSION BQ986943  
VERSION BQ986943.1 GI:22404468  
KEYWORDS EST.  
SOURCE Lactuca sativa  
ORGANISM Lactuca sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;

```

Cichorieae; Lactuca.
1 (bases 1 to 706)
REFERENCE
AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, C., Ellison,
P., Kolkmann, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z.,
Church, S., Jackson, L., and Bradford, K.
Lettuce and Cauliflower ESTs from the Composite Genome Project
http://comgenomics.ucdavis.edu/
JOURNAL
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Amsundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QG_CA_Contig3108, see http://cgdb.ucdavis.edu/
for details.
Plate: QGF10 row: N column: 2.
FEATURES
Location/Qualifiers
1..706
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="L.serriola"
/db_xref="taxon:4236"
/clone="QGFLON21"
/lab_host="E.coli"
/clone_lib="QG_EFGH3 lettuce serriola"
/note="Vector: pBRCDNASFIAB; The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgdb.ucdavis.edu/
TAG LIB=QG_EFGH3 lettuce serriola
TAG_TISSUE=germinating seeds
TAG_SEQ=TCTGTGGGG"
SE COUNT 195 a 110 c 190 g 211 t
IGIN

Alignment Scores:
ed. No.: 0.882 Length: 706
ore: 88.50 Matches: 51
Percent Similarity: 41.36% Conservative: 28
Best Local Similarity: 26.70% Mismatches: 53
Query Match: 8.93% Indels: 59
: 13 Gaps: 13

-09-613-486-15 (1-198) x BQ986943 (1-706)
13 ValLeuThrAspAlaSerLeuAsnGlyValAspLysLeuSerAiaGluVal 32
137 GTTTTACCAGAT---GGGTCCATACCATGATGTGACTCCAGTCTCTCGGTCAG---190
33 GluLysMetLeuValGlnCysGlyAlaProAsnGluGlyIleGluValValPheCysLeu 52
191 -----CTGCACACACGAGGGGCTTCGGGAAGGTGT-----223
53 LeuLeuTyraLaLeuAlaAlaArgThrThrSerProLys-----ValGlnArg 68
224 -----AGGAAGCTGCTCCAAAGATCTTGAGCCGATTATGAGA 262
69 AlaAsp-----SerAspValIlePheSer-----76
263 GTTGAAGTTGTGACGGCTGAGAACATTTGGGGGATGTGATGGGATTTGAATTCGAGA 322
77 -----AsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLys 92
323 AGAGGGCAGATTAATACCTTGTGTGATAG-----CCTGGTGCAGCTT---364

```

93 LysValLeuAspGlyCysAlaProLeuThr-----ArgPheThrAsnLysLeuArg 109  
 365 AAGGTGGTGGATTCCTTTGGTCCACTTGGGAAATGTTTGTAGTATGTGAGTACTTTAAGG 424  
 110 -----ThrPheGlyArg---ThrPheThrGluAlaTyrValAspPheCysIleAlaTy 126  
 425 GGATTCACAAAGAGGAGAGCTCTTACACAATGCAATAGCTAAATTTGATGTTGCGCT 484  
 127 LysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAspSer 146  
 485 CAACATATTCAAGAACAGCTTCTGCTCTAAAGAAAGAGCTGTGGCAGCTTGATCTTCA 544  
 147 -----TysLeuAla-----AlaAspPheLeuGlyThrCysPro 157  
 545 TCTTCATCTTGTGCTATTTAATGATAGGGGAATTTCTGATTTTCATCTTCATATGTCAA 604  
 158 LysLeuSerGlnLeuGlnSerArgLysMet 168  
 605 TATAGGAAAAAACAACAAACAAAGGATT 637

RESULT 2  
 BZ051926/c  
 LOCUS 748 bp DNA linear GSS 09-OCT-2002  
 DEFINITION jcr57h07.g1 B.oleracea001 Brassica oleracea genomic, genomic survey  
 sequence.  
 ACCESSION BZ051926  
 VERSION BZ051926.1 GI:2365291C  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids II; Brassicales; Brassicaceae; Brassica.  
 REFERENCE 1 (bases 1 to 748)  
 AUTHORS Delehaanty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T., Nash  
 ,W., Rabinowicz,P.D. and Wilson,R.K.  
 TITLE Whole genome shotgun reads from Brassica oleracea  
 JOURNAL Unpublished  
 COMMENT Contact: Richard K. Wilson  
 Genome Sequencing Center,  
 Washington University School of Medicine  
 Email: submissions@watson.wustl.edu  
 Plate: jnr57 row: h column: 07  
 Seq primer: -28RFPOT reverse  
 Class: shotgun  
 High quality sequence start: 56  
 High quality sequence stop: 551.  
 FEATURES  
 Location/Qualifiers  
 1..748  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3712"  
 /clone\_lib="B.oleracea001"  
 /note="Vector: pOTw13; whole genome shotgun library from  
 flowering buds. DNA was purified from a crude nuclear  
 prep using Brassica oleracea TOL000DH3 buds provided by  
 Thomas Osborn at the University of Wisconsin.  
 DNA was provided by Pablo Rabinowicz (CSH) and the  
 shotgun library prepared at Washington University Genome  
 Sequencing Center."  
 BASE COUNT 209 a 164 c 136 g 227 t 12 others  
 CRIGIN  
 Alignment Scores:  
 Pred. No.: 0.96 Length: 748  
 Score: 88.50 Matches: 33  
 Percent Similarity: 47.00% Conservative: 14  
 Best Local Similarity: 33.00% Mismatches: 40  
 Query Match: 8.93% Indels: 13  
 DB: 28 Gaps: 5  
 US-09-613-486-15 (1-198) x BZ051926 (1-748)

```

84 vaValvalThrGluGlyAspLeuLysIysValIeu-----AspGlyCysAlaProIeu 131
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
551 ATAACCATCGTGGAAAGGACCGCTAGCAGCAAAATCTTAGGGAGGTGATGGGCTCAGCATGCTC 492
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

102 ThrArgPheThrAsnLysLeuArgThrPheClyArgThrPheThrGluAlaTyrValAsp 121
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
491 ACTTTTAGACTTAAACAAGAGCAGAACATATGCGCGGTATGGTCTCGAAGGTAGCACAC 432
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

122 PheCysIleAlaTyrLys---HisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
431 TTTGATCTCAAGSAGGAAGGTACAAAGATTACAGGGTTCATCGACGAGCTGGCGCGACT 372
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

141 IleProAlaGluAspSerTyrIleuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
371 ATTAGTGCTATTGGAGTTTACTTTAGCTTCCA-----GTAGGCACCATCCCC---TTGACT 321
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGly 182
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
320 CCTGCACAACCAACCAAGAAG-----CTAGAAGCTAAGGGTGGT 282

      SULT 3
      132093
      CUS
      CUS
      DEFINITION
      GENE
      CESSION
      RSION
      YWORDS
      URCE
      ORGANISM

      REFERENCE
      AUTHORS
      TITILE
      JOURNAL
      MMENT

      ATURES
      source
      /organism="Toxoplasma gondii"
      /mol_type="rRNA"
      /strain="RH (Type 1)"
      /db_xref="taxon:581"
      /clone="TGESTzyb06e07.y1"
      /dev_stage="tachyzoite"
      /lab_host="PH108 (GeneHoq, Invitrogen, Inc.)"
      /note="Vector: pBluscript SK-; Site_1: EcoRI; Site_2: XhoI;
      : Toxoplasma RH strain tachyzoites were grown in human
      foreskin fibroblast cultures in vitro. The library was
      originally constructed by K.L.Wan, Cambridge University.
      cDNAs were synthesized from polyA RNAs by oligo d(T)
      priming and directionally cloned into the EcoRI to XhoI
      sites of the Lambda ZapII vector using the ZAP-cDNA
      synthesis kit (Stratagene). The primary cDNA library was
      mass excised as phagemid using EXassist helper phage
      (Stratagene). Phagemid DNA was extracted by
      phenol-chloroform method, and hybridized against a pool of
      over-represented ESTs (N>=12, from 5596 previous reads)".

```

The subtracted library was electroporated into DH10B (GeneHog, Invitrogen, Inc.). WARNING: the library contains a small percentage of cDNAs derived from the human host cells. Library Source: David Sibley, Washington University."

BASE COUNT    1:8 A    1:21 C    1:14 G    1:24 T

ORIGIN

Alignment Scores:

Pred. No.:                 0.776                 Length:                 477  
Score:                      87.00                 Matches:                 30  
Percent Similarity:        43.62%                 Conservative:            11  
Best Local Similarity:     31.91%                 Mismatches:            33  
Query Match:               8.78%                 Indels:                  20  
DB:                          12                         Gaps:                    4

US-09-613-486-15 (1-198) x BM132093 (1-477)

QY    88 GluGlyAspLeuLys---ValLeuaspGlyCysAlaProLeuThrArgPheThrAsn 106  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db    140 GAGGGAGGTGGCGACATAGTTCTTCGACTGTGTCCCCCTTTCAGGTACACGAAC 199  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY    107 LysLeuArgThr-----PheGlyArgThr 114  
::: ||| :::  
Db    200 GAACACAGGATGGCGATACAACATTCTCTCCAGTCTTCATGCTTTTCAGGATTG 259  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY    115 PheThrGluAlaTyValAspPheCysIle-----AaTy:LysHis 128  
||| ::: ||| :::  
Db    260 TTCGCCAGCACCACTGCTATTGTAITACAGTGCTGTGGGCTTCTATTGAAAG 319  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY    129 LysLeuProGlnLeuAsnLaAla---AlaGluLeuGlyLeProAlaGluAsSerTy: 147  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db    320 AAATGCCCGGPGCAATCTGCTTTCCGGAACAATACTGTGACGC"GAGCATCACAGA 379  
||| ::: ||| :::  
QY    148 LeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGlu 161  
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db    380 CAGAGTACATCGACATCGCGCAAAATGTCCGGCATCAGTGAA 421  
||| ::: ||| :::  
  
BQ459495                      562 bp    mRNA           linear    EST 30-MAY-2002  
BA08017r HA Hordeum vulgare subsp. vulgare cDNA clone Ha08017  
5-PRIME, mRNA sequence.  
BQ459495                      GI:21267266  
EST.  
SOURCE                      Hordeum vulgare subsp. vulgare  
ORGANISM                     Hordeum vulgare subsp. vulgare  
Eukaryote; Viridiplantae; Siretophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticeae; Hordeum.  
REFERENCE                    1 (bases 1 to 562)  
AUTHORS                      Radchuk,V., Zhang,H., Weschke,W., Potokina,E. and Wobus,U.  
TITLE                        Barley ESTs from developing seeds  
JOURNAL                      Unpublished  
COMMENT                      Contact: Stejn Nils  
Molecular Markers Group, Department Genbank  
Institute of Plant Genetics and Crop Plant Research (IPK)  
Corrensstr. 3, 06466, Gatersleben, Germany  
Tel: 039482-5522  
Fax: 039482-5595  
Email: stejn@ipk-gatersleben.de  
Insert Length: 562 Std Error: 0.00  
Plate: 8 row: 0 column: 17  
Seq primer: M13rev.  
Location/Qualifiers  
1 . 562  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="barke"  
/db\_xref="taxon:112509"  
/clone="HA08017"  
/tissue type="embryosac"

FEATURES  
source





```

77 -----AsnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLys 92
370 AGAGGGCAGATTAATAGTCTGGGTAGTAAG-----CCTGGGACTT--- 411
93 LysValLeuAspGlyCysAlaProLeuThr-----ArgPheThrAsnLysLeuArg 109
412 AAGGTGGTGATCTTTGGTCCCACTTCGGGAATGTTTCAGTATGTGAGTACTTTAAGG 471
110 -----ThrPheGlyArg-----ThrPheThrGluAlaTyrValAspPheCysIleAlaTyr 126
472 GGAATGACAAAAGGAGAGAGCTCTCTACCAATGCAATTAGCTAAATTTGATGTTGGCCT 531
127 LysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAspSer 145
532 CAACATATTCAAGACGAGCTTTCTGCTGCTANAGAGAAGCTGTGGCAGCTTGATCTTCA 591
147 -----TyrLeuAla-----AlaAspPheLeuGlyThrCys 156
592 TCTTCATCTTTGTGCTATTATATGATAGGGGAAATTTCTGATTTTCATCTTCATATGT 648

RESULT 6
BJ433495/c 671 bp mRNA linear EST 13-MAR-2002
BJ433495 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone ddv22f05 3', mRNA sequence.
BJ433495.1 GI:19408217
EST.
Dictyostelium discoideum
Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
1 (bases 1 to 671)
Urushihara,H.,Tanaka,Y., Kohara,Y. and Shin-i,T.
Full length cDNA of Dictyostelium discoideum at the vegetative
stage
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1..671
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv22f05"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"
19 a 129 c 96 g 227 t

Alignment Scores:
ed. No.: 2 Length: 671
ore: 85.50 Matches: 46
cent Similarity: 35.4% Conservative: 21
st Local Similarity: 24.34% Mismatches: 85
ery Match: 8.63% Indels: 37
3: 12 Gaps: 5

-09-613-486-15 (1-198) x BJ433495 (1-671)
33 GluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIleGluValValPheGlyLeu 52
663 GAGAGGTTAAGTTAGTGGTAAGACACCAATTAGAGATTGAGAGATTGGGAGTATTA 604
53 LeuLeuTyrAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 72

```

```

Db 603 TCATCATCATTTGATTCAGCTCATATGACA-----CAATTAACAAGTGAGAT 556
Qy 73 ValLLePheSerAsnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLys 92
Db 555 TTAGATTTCACAGCAAGAGAGTGGTATTAATGTTGTACATGTCACAGATCAAAATTTAAAG 496
Qy 93 LysValLeuAspGlyCysAlaProLeuThrArgPhe----- 104
Db 495 TTGGGTGTTTAAAGGTATTTCACAGATTTCATAAAGCTTTTAAACAAGGTGTAACGTTTCA 436
Qy 105 -----ThrAsnLysLeuArgThrPheGlyArgThrPheThr 116
Db 435 GTTGGTACTGATCTCCCGCTAGTAGTATGATTTAGATATGTTGGTGAATTACGCTACT 376
Qy 117 GluAlaTyrValAspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAla 136
Db 375 GCAGCTTACATTGATAAACTCTCTGCAAACTACTCATTTCAATT----- 334
Qy 137 AlaGluLeuGlyIleProAlaGluAspSerTyrLeu----- 148
Db 333 ---GAAGGTGGTGAACCTGTAAACCATCATATAGATTTTGTCAATGGCAACTATTAAC 277
Qy 149 AlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSerArgCysMet 168
Db 276 GGTGCAAGGCAATTGGGTATCGATGATAAAGTTGGTTCACCTTCAAAATTCGTAAA----- 223
Qy 169 PheAlaSerMetTyrAlaLeuLysThrGluGlyValValAlaAsnThrProValSerAsn 188
Db 222 TTTCAGAGATTTCATCGCTGTCAAAGTTTCAAGTCATCCAGTCTATGATCCAAATCTCTCAT 163
Qy 189 LeuArgGlnLeuGlyArgArgGluVal 197
Db 162 TTAGTTTATGTTGGTACTAATCATGTC 136

RESULT 7
BZ434184/c 704 bp DNA linear GSS 13-DEC-2002
BONFR91TF BO_1.6_2_KB tot Brassica oleracea genomic clone BONFR91,
genomic survey sequence.
BZ434184
BZ434184.1 GI:26684219
GSS.
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.
Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other_GSSs: BONFR91TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..704
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BONFR91"
/clone_lib="BO_1.6_2_KB tot"
/note="Vector: pBOS1; Site1: BstXI; 1.6-2 kb sheared
total DNA inserted into pBOS1 using BstXI linkers"
169 a 168 c 144 g 223 t

BASE COUNT
ORIGIN

```



Db	333	---	GAAGGTGGTGAACCATCATATATAAGATTTTGTCAATGGCAACTATTAAC	277
Qy	145	AlaAlaAspPheLeuGlyThrCysProLysLeuSerG	uLeuGlnGlnSerArgLysMet	168
Db	276	GGTGCAAGGCATTGGGTATCGATGATAAAGTTGGTTCAC	TTCAAATTGGTAAA-----	223
Qy	169	PheAlaSerMetTyrIleAlaLeuLysThrGluGlyGlyValValAsnThrProValSerAsn	188	
Db	222	TTTGCAGATTTCATCGCTGTCAAAGTTTCAAGTCATCCAGTCTCATGATCATCAATCTCTCAT	163	
Qy	189	LeuArgGlnLeuGlyArgArgGluVal	197	
Db	162	TTAGTTTATGTTGGTACTAATCATGTC	136	
RESULT	10			
SEQ-6284				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
source				
BASE COUNT	201	a	110	c 175 g 168 t
ORIGIN				
Alignment Scores:				
Pred. No.:	2.24			Length: 654
Score:	85.00			Matches: 41

Percent Similarity:	39.78%	Conservative:	31
Very Local Similarity:	22.65%	Mismatches:	89
Every Match:	8.58%	Indels:	20
:	13	Gaps:	6

-09-613-486-15 (1-199) x BQ916284 (1-654)

```

11 AsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLysLeuSerAla 30
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
23 GATCTTCGAGTCGCGAGTAAGCCGGTTATTCGGGTAAATGTCAAA---GGCTCTGT 79
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
31 GluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIleGluValAlaP 50
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
80 TCTGTTAAAAAGCTATGTCGACAATAAGGCGGTGAATACGGAACATAAAGTTAATAC 139
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
51 GlyLeuLeuTyrAlaLeuAlaAlaArg-----ThrThrSer 63
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
140 GATCTTCTTGATGATCGGAGCGCGAATGAGGGACTACTAGACCCTGTTGAGGATCAA 199
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
64 ProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGlyGlu---Arg 82
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
200 GTCCCGATTGTTCGCAGACAGTTCGGCGCGTCTTAGCAGCGGCAGTGGCGGATTGCC 259
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
83 AsnValValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaProLeuThr 102
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
260 GGAATGATGTTGACTGAGGA-----AAAGTAACAAAGATTGCGGAGTT----- 304
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
103 ArgPheThrAsnLysLeuArgThrPheGlyATGThrPheThrGluAlaTyrValAspPhe 122
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
305 -----AGGGTTCAGAGCAAGGAAACACACTTTTGTTCGTCTCTTGATTC 352
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
123 CysLeuAlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaAlaGluLeuGlyIlePro 142
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
353 TTGAGACGGATCAAGGAAGATGTTAAAGAGGTAAACCGCGGACTAGAGTGTGGTATCGG 412
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
143 AlaGluAsp-----SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
413 GTTGATGACTTTAATGATCGGAGGAAGGTGATGTAAATCGAAGCTTTTAAATCCATTCAG 472
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGly 180
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
473 AAAAGCGCAACACTCGAAGGAAGCATCGAGTACCATGACGACGACATTTAAAGAAGTCTGA 532
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

SUBJ 11	BQ048646	610 bp	mRNA	linear	EST 14-AUG-2002
CUS	QGA7M01.YZ.ab1	CG_ABCDI	lettuce	salinas	Lactuca sativa cDNA clone
FINITION	QGA7M01		YV		
CESSION	BQ048646				
RSION	BQ048646.1	Gr:22234115			
YWORDS	EST.				
URCE	Lactuca sativa				
ORGANISM	Lactuca sativa				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae; Lactuca.				
AUTHORS	1 (bases 1 to 610) Kozik A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, B. and Bradford, K.				
TITLE	Lettuce and Sunflower ESTs from the Compositae Genome Project				
JOURNAL	Lettuce and Sunflower ESTs from the Compositae Genome Project				
COMMENT	Unpublished Contact: Alexander Kozik [R.W.Michelmore, Department of Vegetable Crops, R.W.Michelmore Lab University of California at Davis (UCD) Agriculture Hall UCD, Davis, CA 95616, USA				

Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@atgc.org [michelmore@vegmil.ucdavis.edu]  
 belongs to contig QG\_CA\_Contig3108, see <http://cgdb.ucdavis.edu/>  
 for details.

```

FEATURES
  source
    state: OGA, LOW: M, column: 01.
    Location/Qualifiers
      1..610
        /organism="Lactuca sativa"
        /mol_type="mRNA"
        /cullivar="Salinas"
        /db_xref="taxon:4236"
        /clone="QGA7M01"
        /lab_host="E.coli"
        /clone_lib="QG ABCDI lettuce salinas"
        /note="Vector: pBRCDNASFIAB: The library was constructed
        from 10 different sources of RNA from a single genotype.
        Separate cDNAs were generated using primers that
        incorporated unique 5' and 3' tags to distinguish each
        source of RNA. cDNAs were then pooled, size-fractionated,
        directionally cloned into a custom medium-copy vector and
        transformations made with four size classes to minimize
        size bias. Details of each source of RNA and library
        construction can be obtained at http://cgdb.ucdavis.edu/
        TAG LIB-QG ABCDI lettuce salinas
        TAG_TISSUE=shoots environmental stress

```

```

BASE COUNT      157 a      104 c      175 g      174 t
ORIG:N
Alignment Scores:
Pred. No.:      2.35      Length:      610
Score:          84.50     Matches:     50
Percent Similarity: 40.78%  Conservaive: 23
Best Local Similarity: 27.93% Mismatches:  47
Query Match:    1.53%     Indels:     59
DS:             13       Gaps:       13
JS-09-613-486-15 (1-198) x B0848646 (1-610)

```

Qy	13	ValIleThrAspAlaSerSerLeuAsnGlyValAspLysLysLeuLeuSerAlaGluVal	32
Db	137	GTTTTACCCGAT---GGGTCTTACCATGATTTGACTCCAGTCTTCCTTGGTTCCTCAG---190	
Qy	33	GluLysMetLeuValGluLysGlyAlaIleProAsnGluGlyIleGluValValPheGlyLeu	52
Db	191	---CTGCACGACGAGGGGCGTTTCGGGAGGTCTT-----223	
Qy	53	LeuLeuTyrlAlaIleuAlaAlaArgThrThrSerProLys-----ValGlnArg	68
Db	224	---AGGAAGCTGCTCCAAAGATCTTTGGAGCCGATTATGAGA	262
Qy	69	AlaAsp-----SerAspValIlePheSer-----76	
Db	263	GTTGAAGTTGTGACGCCTCAAGACATTTGGGGATGTGATTGGGAGTTTGAATTCGGGA	322
Qy	77	-----AsnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLys	92
Db	323	AGAGGCGACGATTAAATAGCTTTGGTGATAG-----CTCGTGTGGACT---364	
Qy	93	LysValLeuAspGlyCysAlaProLeuThr-----ArgPheThrAsnLysLeuArg	109
Db	365	AAGGTGGTGATCTTTGGTCCCACTTCGGGAAATGTTTCAGTATGTGAGTACTTTAAGG	424
Qy	110	-----ThrPheGlyArg---ThrPheThrGluAlaTyrlValAspPheCysIleAlaTyrl	126
Db	425	GGAATGACAAAAAGGAAGAGCTCTTACACAATGCAATTAGCTAAATTTGATGTTGTGCCT	484
Qy	127	LysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAspSer	146
Db	485	CAACATATTTCAGAACCCAGCTTCTGCTSCCAAGAAAGAAAGCTGTGCAGCTTGATCTTCA	544
Qy	147	-----TyrLeuAla-----AlaAspPheLeuGlyIleThrCys	156





